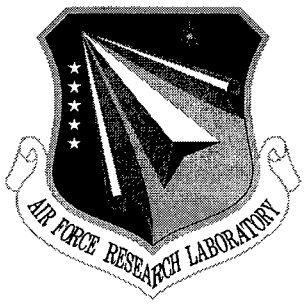
AFRL-SN-RS-TR-1999-201
In House Report
October 1999



REVIEW OF QUANTUM ELECTROMAGNETIC STATES

Michael A. Parker

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AIR FORCE RESEARCH LABORATORY
SENSORS DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK

This report has been reviewed by the Air Force Research Laboratory, Information Directorate, Public Affairs Office (IFOIPA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

AFRL-SN-RS-TR-1999-201 has been reviewed and is approved for publication.

APPROVED:

ANDREW R. PIRICH
Chief, Photonics Processing Branch

FOR THE DIRECTOR:

ROBERT G. POLCE
Chief, Rome Operations Office

If your address has changed or if you wish to be removed from the Air Force Research Laboratory Rome Research Site mailing list, or if the addressee is no longer employed by your organization, please notify AFRL/SNDP, 25 Electronic Pky., Rome, NY 13441-4515. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document require that it be returned.

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 1999	In House, Oct 96 - Oct 98

4. TITLE AND SUBTITLE

REVIEW OF QUANTUM ELECTROMAGNETIC STATES

5. FUNDING NUMBERS

PE - 62702F
PR - 4600
TA - P1
WU- 30

6. AUTHOR(S)

Michael A. Parker

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

AFRL/SNDP
25 Electronic Pky
Rome, NY 13441-4515

8. PERFORMING ORGANIZATION REPORT NUMBER

AFRL-SN-RS-TR-1999-201

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFRL/SNDP
25 Electronic Pky
Rome, NY 13441-4515

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

AFRL-SN-RS-TR-1999-201

11. SUPPLEMENTARY NOTES

AFRL Project Engineer: Michael A. Parker/SNDP/(315)330-4918.

12a. DISTRIBUTION AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This program provided a review of the conceptual and mathematical apparatus to apply modern theories of electromagnetic fields. As a study in low-noise electromagnetic states, it centers on the need to transition basic physics into the engineering arena for improved systems and devices. A survey of publications and texts written during the past 10 years shows that there is a collection of common prerequisite background material; most of this material is addressed in this report. An initial chapter outlines the importance of quantum electromagnetic states to small systems of atoms or photons, to interconnects and links (communications), A-D converters, and to signal processing. All of these systems require higher speed, smaller size, lower power and lower noise components, especially for space applications. The report discusses underlying concepts for Fock, coherent, and squeezed electromagnetic states. It also reviews the density operator, reservoir theory, fluctuation-dissipation theorem and Langevin sources.

14. SUBJECT TERMS

Optical interconnects, Links, Low-noise, Quantum Electromagnetic States

15. NUMBER OF PAGES

300

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

Table of Contents

Table of Figures

Foreword

The purpose of this report is to provide a review of the conceptual and mathematical apparatus to apply modern theories of electromagnetic fields. As a study in low-noise electromagnetic states, it centers on the need to transition basic physics into the engineering arena for improved systems and devices. A survey of publications and texts during the last 10 years shows that there is a collection of common prerequisite background material. This report addresses most of these common areas. Chapter 1 outlines the potential importance of the quantum electromagnetic states to small systems of atoms or photons, to communications systems, and to signal processing. The need for high speed, small size, lower power and low noise components impacts many technical areas including optical interconnects for VLSI, RF links, and signal processing. Although of significant importance, the areas of "thresholdless" lasers, cavity QED, and applications of quantum states are not discussed.

As a special mention, some of the physical pictures presented are the author's own interpretation usually meant to provide intuition. As the reader is certainly aware, it is nearly impossible to provide pictures in quantum mechanics that are entirely consistent with the mathematics; they provide "shadows" of the "real picture". For this reason, some of the figures in the text should be taken as intuitive pictures that do not necessarily embody the full accuracy of the theory.

The chapters are arranged as follows. Chapter 1 contains portions of a proposal indicating some areas of application including RF links, A-D converters and signal processing. Chapter 2 is a review of elementary quantum mechanics and is meant as a convenience to the reader. The first several sections in the chapter are meant to review the relevant notation. The last section introduces the density operator in some detail. Chapter 3 discusses the transition from classical to quantum field theory whereby classical amplitudes are replaced with operators. In the quantum field theory, the modes represent the wave aspects of the field while the operators represent the particle character. Chapter 4 discusses three types of quantum states: Fock, coherent and squeezed states. It briefly introduces methods of producing coherent states (besides lasers) and of measuring the noise in squeezed states. Chapter 5 discusses the master equation (Liouville equation) for the density operator. The first several sections discuss the typical phenomenological master equation. The remaining sections derive the same equation using the fully quantum mechanical procedure. The work clearly shows the origin of the fluctuation-dissipation theorem and the Langevin noise terms. Finally, Chapter 6 shows the use of the density operator to explain the phenomena of laser gain quenching.

Acknowledgments

This technical report originates from two sources. The first source is the author's work with gain-quenched lasers starting in 1989 as collaboration between the AFRL Photonics Center and Cornell University (Prof. C. L. Tang and his group). The work especially grew from the study of noise for a spontaneous emission filter that was first disclosed in 1992 and subsequently patented. The second source consists of graduate courses that the author instructs on semiconductor lasers, boundary value problems, and topics in quantum optics. A large portion of the lasers course covers quantum mechanics, the Liouville equation for the density operator and low-noise states of light.

The author thanks Mr. T. Kemerley as Chief of SND for his suggestion to include the application for low-noise electromagnetic states and nanophotonics into a follow-on technical report. The author also appreciates his patience during the writing of this report. Much thanks to Mr. A. R. Pirich, Chief of the AFRL Photonic Signal Processing branch (SNDP) for his many clear insights on the role of programs and personnel at AFRL. A "tip of the hat" to Sir Paul Cook for many stimulating and often enlightening conversations. Although lasers were not used during the execution of this project, much appreciation to D. Hinkel and T. McEwen for tireless efforts with the argon and TiSa laser for the Device Characterization Facility. Thanks to Dr. D. Hanson for introducing the author to the Israeli photonics program and to JPL, and for his comments on a draft proposal on nanometer scale devices (some of which appears in the first chapter). The author gives a "nod of acknowledgment" to Dr. R. Michalak for early discussions on coherent states. Many thanks to T. Knutti for assistance with the format of the report.

The author thanks Dr. H. Burstyn (patent attorney) and the Jag office at the AFRL Rome Research Site for their continued council and their effort with a number of patents that came to fruition during the execution of the 4600P130 work unit. Two of the patents include semiconductor lasers with intracavity modulators (disclosed in early 1993) and VCSEL quenched lasers.

The author thanks the Electrical Engineering and Computer Science program at Syracuse University for the opportunity to teach (and also to learn) semiconductor lasers (ELE 787) and Boundary Value Problems (ELE 602). This technical report includes portions of the class materials copyrighted under the titles of "The Quantum Laser" (1998), "Boundary Value Problems and Linear Algebra" (1997), and "Quantum Optical Engineering" (1998). Thanks are due especially to Prof. P. Kornreich, Prof. Q. Song and Prof. Arvis for their continued support of the AFRL programs. Much appreciation to my wife, Carol, for her extreme patience during the weekends and evenings over the past several years while I prepared for the courses and compiled the material (and wrote portions of this report). But most of all, I thank the students attending the courses for their challenging questions and answers!

Chapter 1: Introduction

This technical report discusses many of the concepts and mathematical constructs necessary to apply modern theories of electromagnetics. The material comes from many publications and books (some of which are listed in the references). The idea of low-noise electromagnetic (EM) states originated with Glauber (early 1960s) and Yuen. Sometimes the report generically refers to electromagnetic fields as light (quantum states of light), although the concepts apply equally well to radio frequency (RF) or to maser emission (microwave amplification by stimulated emission of radiation).

This technical report is a first part covering common prerequisite material necessary to apply quantum electromagnetic states. A description of low noise EM states is of little value without the possibility of these EM waves interacting with matter. The interaction produces absorption for photodetectors and gain for emission. For this reason, chapters are included on the master equation for the density operator (Liouville equation). The vast majority of analysis uses a direct-product Hilbert space for the matter and EM waves. Almost all of modern laser theory (i.e., matter-field interactions) is stated in terms of the density operator. It is possible to solve the master equations although the first part of this sequence does not do so.

Section 1.1: Comments on the Signal-to-Noise Ratio

The signal-to-noise ratio S/N is important for modern communications systems (including RF links and optical interconnects). As is well known, there are many types of noise (refer to the Cox paper). Shot, thermal and RIN are noise sources of particular importance for photonic systems. Shot noise occurs in emitter-detector systems and origantes in the quantum nature of matter and fields (as this report discusses). Relative intensity noise (RIN) is typically thought of as spontaneous emission, which has the same origins as the shot noise. Referring to spontaneous emission as noise is very specific to the system under consideration; it is not noise for a light emitting diode. There are other sources of noise such as harmonic production by nonlinear devices or mode hopping for lasers. The S/N ration can be improved by either increasing the signal level or decreasing the noise. Increasing the signal level generally requires larger power and physically larger devices. However, larger power and larger devices are contrary to requirements and trends in modern VLSI design and space applications. Therefore reducing the noise level constitutes a desirable alternative approach.

Dynamic range is also well known to be important. It refers to the range of values that an output parameter can assume. If nonlinear behavior is deleterious to the application then the range must be limited to prevent the adverse behavior. For example, a linear transistor amplifier is limited by the onset of saturation as the voltage swings near the supply rails; however, the swing from rail-to-rail can be desirable for digital systems. Obviously, the noise floor limits the dynamic range. By lowering the noise-floor, the dynamic range can be increased.

Noise is potentially more detrimental to analog signals than to digital ones. An analog signal usually caries information on a continuously varying parameter (such as temperature or music) and therefore, the noise determines the ultimate precision of the measurement or the quality of the impressed information. Noise as small as 0.1% can be

significant for audio applications (for example). For a digital system, the impact of noise manifest itself somewhat differently. The digital system is designed with hysteresis and a threshold to provide a clear distinction between a logic "0" and "1". The effects of noise is characterized by the "bit error rate." Of course the distinction is clear to everyone by comparing music from a vinyl record with that from a compact disk.

The electromagnetic field carries noise (shot noise) which is important for RF and photonic systems. Although present at all power levels, these non-classical effects are easiest to discuss for low photon numbers. The power level for which the granularity of the field becomes significant depends on the frequency. High frequency sources (GHz and larger) such as the laser or maser require relatively few photons to give a specific power level as compared with low frequency sources (such as AM radio). Generally, low power levels translate directly to small photon numbers.

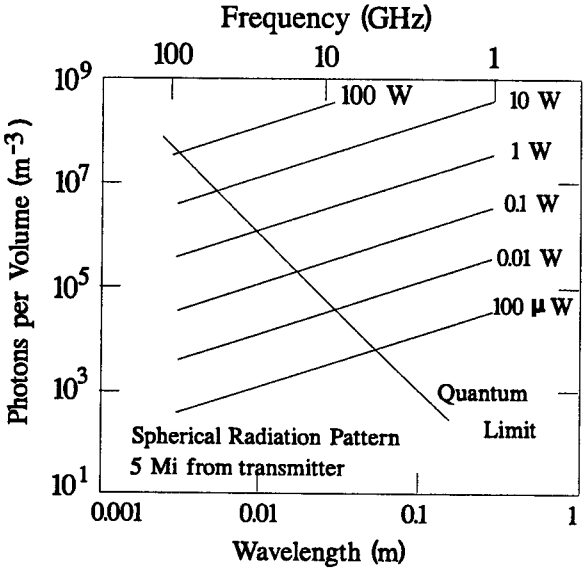


Figure 1.1.1: Graphical illustration of frequency and power of a transmitter for which quantum effects are important.

As mentioned, it is desirable to use low power, small sized devices with high S/N for space applications. At present, there is a modest amount of R&D on systems with small numbers of atoms (nanometer scale devices). Systems with small numbers of atoms are roughly linked with those having a small numbers of photons; a small number of atoms can only emit a small number of photons. Small systems are expected to be more noise-prone (even for EM noise) than larger ones and therefore, the S/N ratio is generally expected to be smaller. For small systems, noise is a problem because small (and low power) components do not deal with many particles (electrons, holes and photons) at one time. For low particle numbers, as might be typical for small or low power components, the uncertainty (or standard deviation) in the signal is roughly the same size as the magnitude of the quantity itself. Equivalently stated, the standard deviation of the number of particles carrying the signal is relatively large compared with the average number.

As another example of noise problems associated with low power systems, consider an RF transmitter or radar. Larger signal-to-noise ratios S/N are one of the reasons transmitters are designed to operate at higher power levels. Figure 1.1 is a plot of the number of photons (per cubic meter) versus the signal wavelength (or frequency – top scale) for specified power levels. The transmitter is assumed to have an approximately spherical radiation pattern and the field is received approximately 5 miles from the transmitter (the photon number is therefore determined at the same distance). The "quantum limit" line represents the collection of points for which the signal level is equal to the inherent quantum noise of the vacuum. For an otherwise perfect receiver and

transmitter, this line gives a signal-to-noise ratio of S/N = 1. Therefore, for conventional electronic equipment (not just optical equipment) operating at modest powers of 100 W and 30 GHz, the quantum noise becomes a factor.

Unfortunately, very few publications apply the basic knowledge of low-noise EM states to engineering problems. The vast majority of published work focuses on the theory and the experiment to verify that theory. The second part of the sequence is aimed at providing a bridge between the requirements of engineering and basic research. Generally, this requires an analysis of the system. The second part of this study includes a comparison of the noise components for systems under consideration as well as engineering applications for low-noise electromagnetic states.

The next few sections provide initial motivation for studying low particle count systems. Although the discussion centers on nanometer scale components, it should be emphasized that other areas (such as RF Links and Optical Interconnects) can be greatly impacted by these small systems.

Section 1.2: Brief Introduction to Quantum EM States

There are several types of noise that are very interesting and highly applicable including spontaneous emission and shot noise. As is well known, atoms can emit light that is coherent with a driving optical field (stimulated emission) and they can also emit light on their own without a driving field (spontaneous emission).

One of the most fundamental notions of Quantum ElectroDynamics (QED) is that of a *Fock* state. For a given spatial volume, Maxwell's ElectroMagnetic (EM) equations can be solved so as to define a set of allowed EM modes (for a cubic volume, the modes are sines and cosines). For Fock states, QED keeps track of the number of photons in a mode, which is characterized by the allowed wavelengths and polarization. A Fock state is a mode with a definite number of photons (this means that each mode has a definite average power). In *classical* electrodynamics, a state without any photons corresponds to a mode without any amplitude. In QED, a state without any photons (the vacuum state) has an *average* electric field of zero, but the variance (which is proportional to the square of the field) is non-zero. This means that the value of the electric field can fluctuate away from the average of zero. As is well known, the non-zero variance refers to quantum vacuum fluctuations or noise; the vacuum state has the minimum quantum noise. There is a slight complication for engineering purposes however: the average electric field is zero for all Fock states and in general, the noise is greater than the minimum value set by the vacuum. In addition, although the exact number of photons is known, the phase of the collection is not known.

A *coherent* state is a state with a nonzero average electric field and a fairly well defined phase. An electric field in a coherent state can be pictured as sine and cosine waves (as typical for classical electromagnetics). Laser emission is best described by coherent states. The coherent state is actually a linear combination of all Fock states (the two types of states are seen to be quite different). One of the most important distinctions is that, for a coherent state with a given (average) amplitude, the number of photons in the mode is described by a Poisson probability distribution. The average number of particles is $\langle n \rangle$ while the standard deviation $\sigma = \sqrt{\langle n \rangle}$. For example, a beam with an

average of $\langle n \rangle$=100 photons, will have a standard deviation of $\sigma = \sqrt{\langle n \rangle}$=10 photons. One might reasonably expect the number of photons to range from 80 to 120 (almost 50% variation). This means that the number of photons in a beam can fluctuate such that the variance of the number is approximately equal to the average number of photons in the beam (shot noise). Now returning to the amplitude and phase, it just so happens that the noise in the amplitude and phase of the coherent state is no worse than that of the vacuum state (regardless of the amplitude of the coherent state). The "Quantum Limit" line in the figure on page 15 represents the vacuum level of noise, and also therefore, the noise level of a general nonzero coherent state. The time dependent electric field can be decomposed into a time dependent sine and cosine term so that the uncertainty in the amplitudes of these quadrature terms can be specified as an alternative to specifying the uncertainty in the amplitude and phase.

A *squeezed* vacuum state is obtained from the quantum vacuum state by reducing the noise in one set of parameters while adding it to another (i.e., "squeezing the noise out"). Squeezing the vacuum is equivalent to squeezing the coherent state since the vacuum and coherent states have the same type and amount of noise. For example, noise can be removed from one quadrature term of the electric field (quadrature squeezing) but that removed noise reappears in the other quadrature term. Similarly, a "quiet" photon stream (i.e. a number squeezed state) obtains when noise is removed from the photon-number but reappears in the phase; a quite photon stream is describe by "sub-Poisson" statistics. Number-squeezed electromagnetic waves are pictured as photons that are equally spaced travelling in the beam. Perfect number squeezed states are Fock states. There are also phase squeezed states. Squeezed coherent states can be produced, detected and used for low noise applications.

Spontaneous emission is another form of noise in the laser. As is well known, the spontaneous emission is necessary to initiate the laser oscillation; however, the threshold power is larger than necessary and produces wasted energy. An interesting fact about spontaneous emission is that it is a result of quantum vacuum fluctuations and that it is not a property of the emitting collection of atoms. The spontaneous emission is initiated by the fluctuating electric field of the vacuum states. The rate of spontaneous emission can be modified by (1) changing the number of vacuum modes (there is one vacuum mode for each wavelength and polarization allowed by the boundary conditions on the enclosed volume) or by (2) shifting the atomic resonance away from the frequency that characterizes the vacuum mode. The relatively new field of Cavity QED describes the theory and measurement of both spontaneous and stimulated emission for which unusual cavity effects become important.

Section 1.3: Links and Interconnects

Links and interconnects are the subsystems that pass a signal from one point to another. There needs to be a transmitter, receiver and waveguiding medium (electronic or optical). The actual configuration and operating characteristics can vary depending on the system employing the link. The link can be either analogue or digital. It might operate at any number of frequencies and bandwidth. Generally, it is desirable to dissipate as little power as necessary while maintaining low noise and high dynamic range. Gain is good. Although most of the links are digital at this point (digital phones, computer backplanes...), some groups continue to examine the properties of the analogue system.

In the early 1990s, the author and members of his group, in cooperation with Cornell University and Syracuse University, developed a number of interconnect devices capable of both linear and digital operation. The set of devices include lasers with built-in modulators, integrated lasers and fets (initiated by Dr. D. Honey), optical-optical gain quenched lasers (refer to the next section), and evanescently coupled photodetectors.

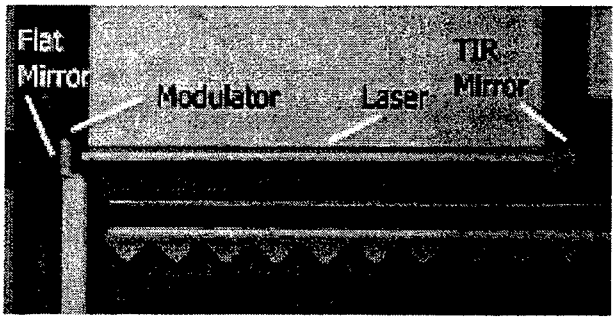


Figure 1.3.1: A laser with a built-in modulator

An example of a semiconductor laser with built-in modulator appears in Figure 1.3.1. Reverse biasing the modulator reduces the net gain in the laser and reduces the power in the output beam. The reverse bias requires as little as a few microamps (at 1.5 volts) to yield a power gain on the order of 10-100. The on/off contrast ratio is better than 250:1. The device can be operated in either a linear or digital mode. For the digital mode, hysteresis can be added by including feedback from the output to the input. The modulator current can be provided by a micron-sized FET, resistor or photodetector. The modulator has a size on the order of 5x5 microns which is about a factor of 50 smaller than an "external" integrated modulator. The modulator can be modeled as a lumped element whereas the external modulator forms a transmission line at GHz frequencies.

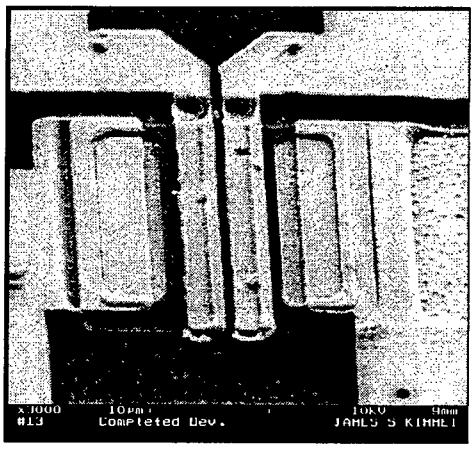


Figure 1.3.2: Dual gate FET monolithically integrated with a laser with built-in modulator (not shown).

Monolithically integrated FETs and lasers can be used to form a link (applied as part of J. S. Kimmet's MSEE thesis). Figure 1.3.2 shows an example of a dual gate FET used to drive the laser. The directly modulated laser is conceptually the simplest method to obtain modulated light. Generally, the semiconductor laser requires large bias currents (for good S/N); this power

requirement translates to the physical size of the FET being substantially larger than the laser.

Numerous systems require interconnects and links. For example, the AFRL sensors directorate is presently interested in space-based-radar (SBR). One project in support of SBR, proposes to develop a high speed optical A-D converter. Among other issues, the system requires a stable pulsed laser system; the laser and other components are preferably monolithically integrated. Another project consists of developing a link operating between the antenna element and the digital transmitter/receiver. Although present trend is to increasing use a digital format for the signal, one version of the link use analog components. The link requires a low noise laser source, fiber and photodetector. An alternate possibility however, is to use a low-noise semiconductor laser (as a local oscillator) and mix (heterodyne) the output with the radar signal at the antenna. A semiconductor laser can be modulated at the lower frequency without many of the noise problems. However this second scheme still requires a low-noise stable local oscillator. For all three of these projects, one key component is a low-noise semiconductor laser operating in either a pulsed or continuous-wave mode. Again, space applications also require low power dissipation since the power generators are expected to have limited capacity and there is no convenient source of cooling.

As discussed by Cox, there are three primary sources of noise for semiconductor laser systems. The three sources of noise are shot, RIN and thermal. Shot noise is primarily associated with the photon statistics of the incident electromagentic wave. RIN is related to the spontaneous emission from a laser. Thermal noise can be controlled to some extent by using wide-bandgap materials and coolers. One approach to developing a low-noise laser system is to apply low-noise electromagnetic states to the problem. The shot noise and RIN are actually interrelated for the low-noise laser. Unfortunately, a full analysis for the application of low-noise EM states is beyond the stated scope of this report. However, it is possible to indicate some guiding principles.

It is possible to design and construct a laser link operating below the shot noise limit. Several recent publications discuss operating a semiconductor laser with 96% of the amplitude noise (i.e., photon number noise) removed from the laser beam. For this level of noise reduction, it is necessary to control the losses in the system including spontaneous emission.

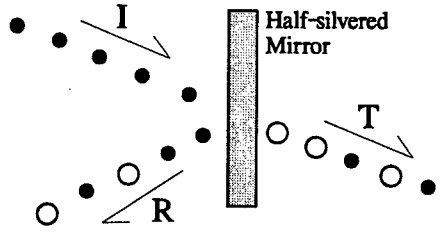


Figure 1.3.3: Photons in incident beam I are 50% likely to be transmitted T or reflected R.

Loss in an optical system is detrimental to low noise EM states. Consider a low-noise number state as might be pictured as a regular stream of photons as shown in Figure 1.3.3. Suppose that this stream is incident on a 50% silvered mirror. As shown in the figure, there is a 50% chance that a photon is detected behind the mirror. The reflected and transmitted beams are therefore randomized and obey binary statistics. The variance of the photon number for the incident beam is zero but becomes nonzero for both the reflected and transmitted beams. The "partition" noise induced by the half-silvered mirror essentially randomizes the incident beam.

It is reasonable to expect that a laser's mirror effects the variance of the photon number. The reflectivity generally needs to be close to either "0" or "1" to prevent

randomizing the photon stream. Fibers need to be well coupled to the output facet. Number squeezing as high as 96% has been observed.

Absorption has similar effect on the probability distribution of a photon stream. For example, a chunk of material that absorbs 50 % of the incident light randomly removes 50% of an incident photon stream. Spontaneous emission from a laser also tends to destroy a low-noise state. One can expect low RIN and squeezed states to be linked.

Methods to lower the spontaneous emission are presently of interest for low-threshold lasers. Spontaneous emission is necessary to initiate laser oscillation. However, any additional spontaneous emission is wasted power. One method to reduce spontaneous emission is to use oxide confinement to better overlap the laser optical mode with the pumped region of the gain medium.

Parametric amplifiers, semiconductor lasers driven by constant current sources, semiconductor lasers with special construction, and devices employing four-wave mixing can produce low-noise EM states (sqeezed states). Homodyne detection systems can be used to measure the noise.

Section 1.4: Some Initial Work on Electromagnetic Noise

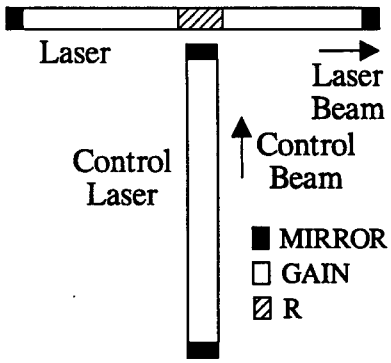


Figure 1.4.1: Gain-Quenched Laser

In 1989, the AFRL Photonics Center began collaborative work with Cornell University (Prof. C. L. Tang and his group) on semiconductor laser devices. Most of the devices are based on gain-quenching where a "control" beam from one laser competes for the gain of another laser and thereby reduces the emitted power for the second laser (refer to Figure 1.4.1). Of course, the most prominent mechanism for the operation of this type of device is the matter-light interaction. Inevitably, some of the control lasers are operated below threshold. In such a case, any change in the output beam of the laser is attributed to spontaneous emission. It is found that the control beam from a control laser below threshold has little effect on the laser beam. A spontaneous emission filter can be made by replacing the control laser with a laser amplifier or a fiber that carries the noisy signal. As discussed in Chapter 7, there are a number of possible reasons that the spontaneous emission does not much influence the output of the main laser. One speculation concerns the efficiency of spontaneous emission to quench the laser. The discussion below considers one scenario whereby the phase of an incident EM wave fluctuates rapidly and inhibits the process of stimulated emission; as a result, the effect of quenching is minimal. In this way, it might be speculated that the "incoherent" portion of the incident light beam does not quench whereas the coherent portion does quench.

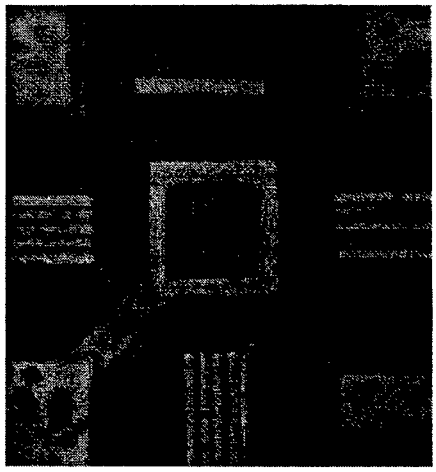


Figure 1.4.2: Queched VCSEL with 3 linear waveguides. VCSEL has ring electrode with output facet at the center. The VCSEL uses oxide confinement.

The following discussion gives an intuitive picture that that leads to the study of low-noise EM states. The same argument can be presented using a density matrix approach. The model is oversimplified since (1) rapid phase variations can render the definition of the classical EM field meaningless, (2) external damping effects should be included, and (3) neighboring electronic states should be included (Fermi's golden rule).

According to the semiclassical theory, the oscillating dipoles must be in phase with the driving EM field for a period of time before a transition can occur (i.e. this time interval is essentially the reciprocal of Rabbi frequency). Now if the phase $\phi(t)$ of the driving field $e^{i\omega t+i\phi(t)}$ changes fast enough, the driving field is not in-phase with the oscillating dipole for a long enough period of time to induce a transition. This is true for absorption or stimulated emission. Various probability distributions for ϕ are possible and lead to various "gains" for the transition. The reader not interested in the mathematical detail should jump to the discussion after Equation 1.4.2.

To illustrate how the "gain of transition" depends on the entropy of the phase, consider the following argument based on time dependent perturbation theory from quantum mechanics. Suppose an atom is perturbed by an incident classical electric field E(t). Assume H_o is the Hamiltonian of the unperturbed atom and

$$\hat{V}(t) = \hat{\mu}E(t) = \hat{\mu}E_o e^{j\omega t + j\phi(t)} = \hat{V}e^{j\omega t + j\phi(t)}$$

where $\hat{\mu}$ is the dipole moment operator, $j = \sqrt{-1}$, and the conjugate term is dropped for convenience (the form of the perturbing potential assumes an emission event). For example, Figure 1.4.2 shows an electric field with multiple discontinuities in the phase. The middle portion of the figure shows the approximate phase as a function of time. As usual with time perturbation theory, assume that $\{|n\rangle = |u_n\rangle\}$ are the complete set of eigenfunctions for the unperturbed hamiltonian H_o

$$\hat{H}_o|n\rangle = E_n|n\rangle$$

The wavefunctions that satisfy Schrodinger's equation for the total hamiltonian

$$\hat{H}|\psi(t)\rangle = \left[\hat{H}_o + \hat{V}(t)\right]|\psi(t)\rangle = j\hbar\frac{\partial}{\partial t}|\psi(t)\rangle$$

are given by

$$|\psi(t)\rangle = \sum_n \beta_n(t)|n\rangle$$

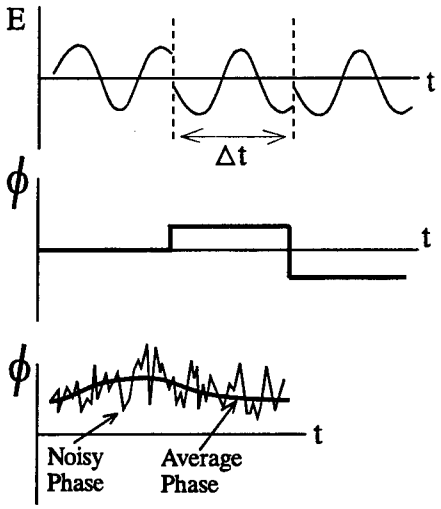


Figure 1.4.3: Top: electric field with phase shift. Middle: Phase of electric field. Bottom: the average of a rapidly varying phase.

Assume that the atom is initially known to be in state "i" so that

$$\beta_n(0) = \begin{cases} 1 & n = i \\ 0 & n \neq i \end{cases}$$

The probability of transition from state "i" to state "f" $(i \neq f)$ is

$$P_{i\to f} = |\beta_f|^2$$

where, to first order perturbation,

$$\beta_f = \frac{1}{j\hbar}e^{-i\omega_f t}\int_0^t d\tau\, e^{j\omega_{fi}\tau}\langle f|\hat{V}(\tau)|i\rangle$$

and

$$\omega_f = \frac{E_f}{\hbar} \qquad \omega_{fi} = \omega_f - \omega_i < 0$$

Substituting the perturbing potential for the matrix element provides

$$\beta_f = \frac{V_{fi}}{j\hbar}e^{-j\omega_f t}\int_0^t d\tau\, e^{j\omega_{fi}\tau}e^{j\omega\tau + j\phi(\tau)} = \frac{V_{fi}}{j\hbar}e^{-j\omega_f t}\int_0^t d\tau\, e^{j\Omega\tau}e^{j\phi(\tau)} \qquad (1.4.1)$$

where, for simplicity, $\Omega = \omega + \omega_{fi}$ is also substituted. Notice that Ω can be quite small and the time scale $T \sim \frac{1}{\Omega}$ can be quite large. *Assume that the time interval over which*

the phase changes is small compared with T. At this point, the integral can be evaluated. For noise impressed on an average varying phase (see bottom portion of Figure 1.4.2) such as $\phi(t)=\bar{\phi}(t)+\phi_n(t)$ where ϕ_n is an ergodic noise process, an average of Equation 1.4.1 can be found using an ensemble of systems. However, for illustration purposes, the integral in equation 1.3.1 is simplified with the help of some simple ideas.

Lemma: Let "F(t)" and "S(t)" be functions (of time) with the properties that "F" is a rapidly varying function and "S" is a slowly varying function. Let the $\hat{\tau}_i$ be a time interval large compared with the duration of the variations of "F" but small compared with the duration of the variation of "S". Let τ_i be a time in the interval $\hat{\tau}_i$. The intervals $\hat{\tau}_i$ might as well all have the same duration τ so that the relation $t=\sum_{i=1}^{N}\hat{\tau}_i=N\tau$ holds. Let $\langle\ \rangle_\tau$ denote an average over the small time interval. The following relations hold

1. If $\langle F\rangle_\tau=0$ then $\langle F\rangle_{t>>\tau}=\frac{1}{t}\int_0^t dt\,F=\frac{1}{N}\sum_n\frac{1}{\tau}\int_{\tau_n}dt\,F=\frac{1}{N}\sum_n\langle F\rangle_\tau=0$

2. If $\langle F\rangle_\tau=0$ then $\langle F(t)+S(t)\rangle_\tau=\langle F\rangle_\tau+\langle S\rangle_\tau=\langle S\rangle_\tau=S(t)$

3. Let Δt_i be small compared with τ so that the definition of Riemann integral holds

$$\int dt\,SF\cong\sum_i S(t_i)F(t_i)\Delta t_i=\sum_{\hat{\tau}_i}\sum_{t_i\in\hat{\tau}_i}S(t_i)F(t_i)\Delta t_i$$

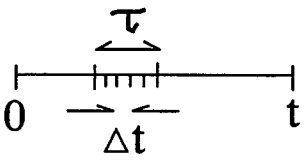


Figure 1.4.3: Comparison of intervals

Since t_i is in the interval $\hat{\tau}_i$ and "S" is slowly varying, the relation $S(t_i)=S(\tau_i)$ holds. The integral can be written as

$$\int dt\,SF\cong\sum_{\tau_i}S(\tau_i)\sum_{t_i\in\hat{\tau}_i}F(t_i)\Delta t_i\cong\sum_{\tau_i}S(\tau_i)\hat{\tau}_i\frac{1}{\hat{\tau}_i}\sum_{t_i\in\hat{\tau}_i}F(t_i)\Delta t_i\cong\sum_{\tau_i}\hat{\tau}_i\,S(\tau_i)\langle F\rangle_{\tau_i}\cong\int d\tau\,S(\tau)\langle F\rangle_\tau$$

where, in the last integral, τ is a dummy variable and the average of "F" is written as a function of time with the subscript. If the average of "F" over the small interval is zero, then the entire integral is zero.
--- End Lemma ---

Returning to Equation 1.4.1, which is

$$\beta_f=\frac{V_{fi}}{j\hbar}e^{-j\omega_f t}\int_0^t d\tau\,e^{j\Omega\tau}e^{j\phi(\tau)}$$

and substituting $\phi(t)=\bar{\phi}(t)+\phi_n(t)$, the expansion coefficient becomes

$$\beta_f=\frac{V_{fi}}{j\hbar}e^{-j\omega_f t}\int_0^t d\tau\,e^{j[\Omega+\bar{\phi}(\tau)]\tau}e^{j\phi_n(\tau)}$$

For illustration, it is sufficient to assume the average phase is zero. Also the exponential term $e^{j\Omega\tau}$ is the slowly varying function while $e^{j\phi_n(\tau)}$ is assumed to be the rapidly varying one. The third part of the Lemma can be used to write the expansion coefficient as

$$\beta_f = \frac{V_{fi}}{j\hbar} e^{-j\omega_f t} \int_0 d\tau e^{j\Omega\tau} \left\langle e^{j\phi_n} \right\rangle_\tau \tag{1.4.2}$$

If there is no phase noise $\phi_n(t) = 0$ for all times, then the average under the integral is unity and the expansion coeffient is

$$\beta_f = \frac{V_{fi}}{j\hbar} e^{-j\omega_f t} \int_0 d\tau e^{j\Omega\tau}$$

which is the usual result from perturbation theory. If the phase varies rapidly enough then the expansion coefficient is smaller than usual. For example, if $\phi = 0, \pi, 0, \pi, \ldots$ then the average of $e^{j\phi_n} = 1, -1, 1, -1, \ldots$ is zero. The probability of transition is zero for this last example.

The discussion above illustrates that sufficiently rapid phase variations (such as for FM modulation or laser phase noise) can inhibit transitions. For such a case, photodetectors might not be expected to absorb light and some laser amplifiers might not amplify. The above reasoning suggests that certain normally opaque materials might be transparent for sufficiently rapid phase modulation. In such a case, if the normally opaque material covers a truely opaque one, the one might speculate that phase modulation might allow a person to view the underlying material without interference from the top one.

The requirement for the simple model presented above is that the driving field be near resonance. The time scale of interest is

$$T << \frac{1}{\Omega} = \frac{1}{\omega + \omega_{fi}}$$

where ω_{fi} is negative since we assume the emission case. Suppose that the bandwidth of the typical quantum well laser diode is used to calculate T. The typical order of magnitude of the bandwidth of spontaneous emission is 5 nm. This is essential the "natural line width" set by the dipole dephasing time. For GaAs semiconductors, the dipole dephasing time is 10 to 100 femtoseconds. The oscillation period of 800 nm light is 3 femtoseconds.

Section 1.5: Comments on Space Applications of Signal Processing

The Joint Warfighting Science and Technology Plan specifies Joint Warfighting Capability Objectives (JWCO) and Key Technologies that are required to meet the JWCO. The JWCO clearly indicates the need for higher performance links and signal processing, greater storage capacity, better components and miniaturization technologies. Space applications, in particular, require low power, small size, light weight, high speed, environmentally rugged components. These Systems use sensors, A-D converters, signal processors, data storage and memory, and rf/optical links.

A typical space application might be a satellite. The variety of sensors and detection methods includes radar, photodetectors (high power or sensitive), detector arrays capable of seeing through clouds, sensors to image subterranean objects, low light TV, high resolution detector arrays for imaging from space platforms, vibration and motion sensors. The EM-wave sensors preferably operate at multiple wavelengths. An A-D converter is envisioned to transform analog sensory data into a digital format for further signal processing, encryption, or RF digital transmission. A space platform requires a signal processor for multiple systems and subsystems such as for space based radar, digital receivers and communications, antenna control, platform maneuvering, data reduction and fusion, data storage for delayed transmission, downlinks and subsystem programming. RF down links can used to transmit compressed data or to send full-length records to a ground station. Optical links are potentially important for communications between multiple space platforms (internet-in-the-sky) since free-space optical beams are difficult to intercept and do not require massive structures to produce them. Lightweight fiber might eventually replace heavier cables on the platforms.

The signal processor in this example scenario is envisioned to consist of photonic integrated circuits, optical interconnects and integrated memory that conform to an overall architecture. The opto-electronic components are monolithically integrated. The portion of the interconnect that caries the optical signal can be free space, fiber, monolithically integrated waveguides or other waveguiding optics between circuits. At present, there is widespread belief that optical interconnects are most important for long-haul transmission (on the order of kilometers) or between gobal systems rather than between devices on small integrated circuits. The integrated circuit is envisioned to consist of detectors, emitters, memory and logic devices.

A realistic program consists of designing and implementing a high performance RF signal processor with a ten-fold improvement in size, weight and power requirements over presently available processors. Such a processor is envisioned to provide two processing and storage functions: (1) store massive amounts of raw data in a down-link buffer and (2) provide real time sensor fusion and image analysis for immediate transmission to a ground station.

The evolution of components for signal processing is toward smaller, faster and cheaper. As stated by JWCO, the AF depends on rugged, fast, light weight components for superior warfighter space-air systems. The performance of present systems is orders of magnitude away from the performance of systems using opto-electronically enhanced modules, memory, processing elements and interconnects.

A program can be aimed at developing small, low power, integrated devices for RF digital receivers, signal processors, and communications equipment. The desired

trend is away from large power hungry multichip modules and toward circuits with the optics and electronics monolithically integrated on a single wafer. Most chip manufacturers agree that there needs to be further decreases in size and power to meet the future needs of industry and DOD. These requirements pose significant problems for both the design and fabrication of the components.

Monolithic integration of opto-electronic components (especially those with nanometer scale features) specifically addresses the size, weight and speed issues of electronic and optical components for systems (as stated by JWCO and TPITs). The linear feature sizes can be a factor of 20 (or more) smaller than presently available devices. Smaller feature sizes directly translate to lower power and higher speed. Small, low power devices or signals require new techniques to handle noise problems. These components can be designed as integrated memory elements with more than a factor of 20x20=400 (per surface area) increase in density over present state of the art integrated memory.

Integrated emitters (lasers/leds) and detectors are appropriate for optical interconnects and links that require high packing density and low power. Low power and small sized devices generally can be expected to exhibit lower signal-to-noise S/N ratios than the larger, more powerful ones. Low power components tend to operate near the noise floor. It is therefore important to identify the noise sources and reduce their effect where possible. The components must be monolithically integrated into a semiconductor wafer. The circuits need to be rad-hard, rugged, lightweight and small for space applications.

RF and digital receivers necessary for the space applications consist of integrated components and signal processors. As mentioned, the receivers can connect to sensors through AD converters. These receivers require low noise, lightweight, low power, small and rugged components. At present, the bottleneck in commercial processing systems is the bus-speed that can be increased by using opto-electronic components and interconnects.

Section 1.6: Discussion of Device Trends

The goal is to improve system performance by investigating and developing low power, small, lightweight, photonic components and subsystems. Small size naturally leads to higher speed by reducing signal propagation and interaction times. Modern research investigates physical systems consisting of a small number of particles ($<10^6$); these "systems" can be either a device or a signal. The term "particles" refers to atoms in a device and photons in a signal. Small signals naturally have poor signal to noise ratios and they have lower dynamic range which can dramatically affect both analog and digital devices. Small devices (scale sizes less than 1000 angstroms) tend to produce relatively large amounts of thermal noise (compared with the allowable signal sizes). One way to improve the statistics is to increase the number of particles in the system. Unfortunately, such systems necessarily involve large sizes and high power. The noise for the entire system must be examined.

Present commercial components have minimum sizes on the order of 200 nm for DRAM, 3x30 μm^2 for in-plane (edge-emitting) lasers, 10x10 μm^2 for VCSELS (with thresholds as low as 0.2 mA), 200 nm gate lengths for FETs, and 1000 nm pixel sizes for CD ROM. Nanophotonic components have features smaller than 100 nm which corresponds to roughly 10^6 atoms or less. Fewer atoms implies smaller signals, lower power dissipation and higher speeds. Low particle-count physical systems require non-classical descriptions since (1) small volumes tend to emphasize quantization effects (electron wavefunctions in first quantization), (2) the collective properties of a small number of atoms or photons are not adequately described by ensemble averages. The departure from the ensemble average (i.e., noise), as represented by a probability distribution, becomes significant for small physical systems that are best described by a form of Quantum ElectroDynamics (QED). However, QED -- the best theory available for electromagnetic interactions -- is a rather general theory and has not been extensively applied to engineering problems.

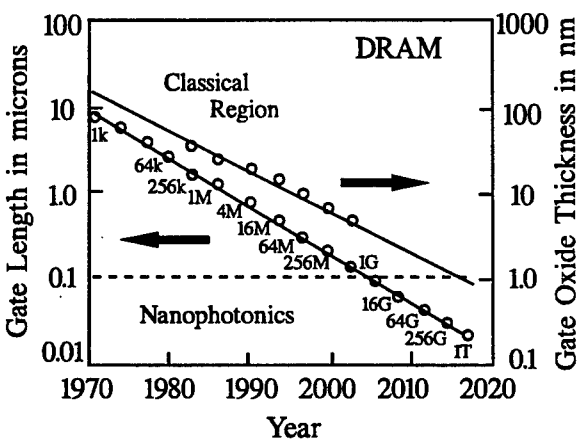


Figure 1.6.0: Device trends.

Besides new applications of physical theory to engineering practice, the fabrication of nanometer scale circuits and devices poses technical challenges. It is relatively easy to grow layers as thin as one or two atoms into a material by using Molecular Beam Epitaxy. However, it is difficult to laterally pattern structures of similar sizes. One important challenge for the commercial sector consists of developing a reliable, low cost lithography (or other means) capable of producing lateral feature sizes smaller than 100 nm. Recently, an ARPA sponsored program has developed an optical photoresist suitable for 80 nm feature sizes. At present, electron lithography is capable of feature sizes as small as 10 nm although the process is expensive. Generally, smaller devices require material with smaller and fewer defects than for presently available

commercial materials. As a side comment, either the number of process steps must decrease or the reliability of each step must increase since the probability of at least one defective component on a wafer increases with the number of components. To minimize the number of wasted-die, manufacturers of integrated circuits use extensive process development and calibration; they do not flippantly change even the smallest of details or process steps. As is evident, the development of new components places new requirements on design, materials and fabrication.

To illustrate the advantages of nanometer scale components, consider the signal processor for a surveillance system. The processor and links consist monolithically integrated and bulk storage memory, optical interconnects and integrated nanometer scale opto-electronics operating at RF speeds. The highest possible speed, using nanometer-scale components is approximately 100 THz with an ultimate packing density of approximately 10 TeraBits/cm^2. This ultimate speed and packing density is based on the speed of light between components that have atomic dimensions. Fewer atoms implies smaller signals, lower power dissipation and higher speeds. The maximum far-term device and processor specifications are given in the next list.

Section 1.7: Brief Review of Some Nanometer Scale Devices

A number of quantum devices have been developed by several internationally known universities and large corporations (refer to Turton's book on Quantum Dots) using either an evolutionary or a revolutionary approach. The evolutionary approach emphasizes scaling down present devices with only minor modifications to the construction and operating principles. The revolutionary approach develops new devices with new operating principles; these new devices do not resemble former ones with respect to design and construction. The previously developed devices provide a convenient starting point for the work on nanophotonic components and can be incorporated into a demonstration of the technology. The Resonant Tunneling Diode (RTD) and Resonant Tunneling Transistor (RTT) both use a "quantum size" layer positioned between two tunneling barriers. With the proper source-to-drain or gate voltage, carriers can tunnel through the barriers to produce current. The Single Electron Transistor (SET) is somewhat similar to the RTT except that only one electron can enter the gate region at one time (due to coulomb repulsion). The Quantum Dot (QD) has two free electrons within its volume; the configuration of these two electrons determines the state of the device. Aharanov-Bohm type devices use a magnetic field to change the phase of the electron wavefunction so that the wavefunction either constructively or destructively interferes with itself so as to produce "on" or "off" states, respectively.

Resonant Tunnel Device (RTD)

A resonant tunneling device has two terminals and it is similar in construction to a normal diode. The device consists of a quantum well region separated from two conduction regions by quantum barriers (on the order of 50 to 100 angstroms thick). The barriers can be AlGaAs, for example, while the well and conduction regions can be GaAs. Conduction through the device occurs by quantum tunneling through the barriers. The voltage must be adjusted across the device so that the energy levels in the well have the same energy as the conduction band on the electron injection side (left side of the figure).

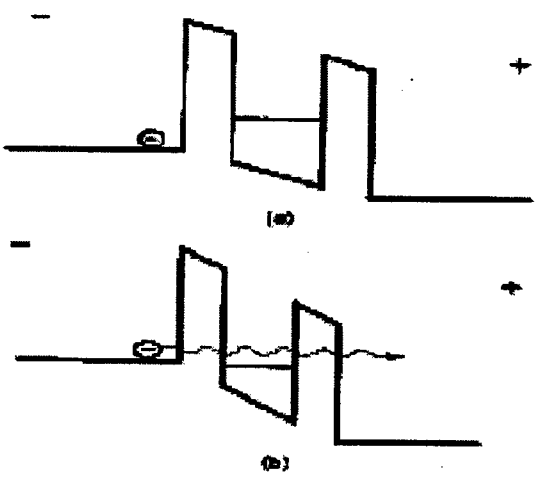


Figure 1.6.1: Band-edge diagram for electron transport in an RTD

Resonant Tunneling Transistor (RTT)

A Resonant Tunneling Transistor is similar in construction to the resonant tunneling device except, in addition to a source and drain, it has a gate electrode attached to the well region. Electrons are injected into the device through the source electrode. An applied gate voltage adjusts the energy of the well levels with respect to the energy range of the occupied conduction states in the source. Carriers tunnel through the barrier into

the well from the source when the energy of a well state falls within the energy range of the occupied conduction states in the source. With a voltage applied from the source to drain, the charge will then flow into the drain.

Single Electron Transistors

A Single Electron Transistor (SET) has gate, drain and source terminals. With appropriate gate voltage, a single electron from the source can tunnel onto a center island (somewhat similar in structure to the RTT). The additional charge changes the energy levels of the island and coulomb repulsion prevents additional electrons from entering the island (for fixed gate voltage). Once the extra electron transfers to the drain, additional charge can enter the island and thereby contribute to the conduction process. The current can be quite large (tens of microamps).

Quantum Dots

Quantum Dots consist of a sequence of islands with each one spatially confining two charges. A length of end-to-end quantum dots can be used to transfer a signal. Charges in neighboring cells assume the same configuration as the "input" (or control) cell, which is shown furthest to the left in the figure. If the charges in the control cell are rotated by 90 degrees then the neighboring cell will assume the same state to minimize the coulomb interaction energy. Subsequent cells will also rotate their state by 90 degrees. The figure shows an inverter as an example. The cell on the right assumes the opposite state from the others due to coulomb repulsion from the neighboring cells.

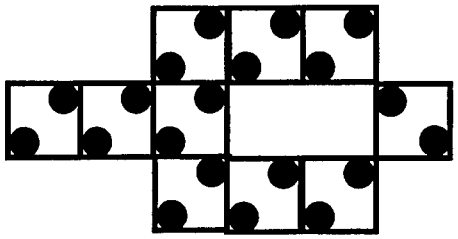


Figure 1.6.2: A quantum dot inverter. The signal propagates left to right.

Aharanov-Bohm Effect Device

The Aharanov-Bohm Effect Device (ABED) is based on the fact that the phase of an electron wavefunction can be influence by the presence of an electric or magnetic field *even when that field is completely isolated from the charge.* The field is assumed to be confined to region A in the figure. By changing the phase of the wavefunction in the top branch vs that for the bottom branch, it is possible to have either constructive or destructive interference which in-turn controls whether current flows or not.

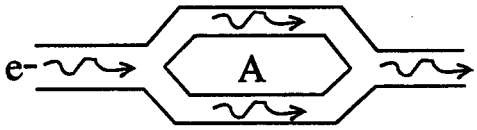


Figure 1.6.3: The Aharanov-Bohm effect device.

Quantum Interference Devices

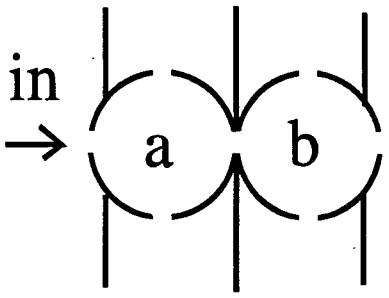


Figure 1.6.4: A representation of the quantum interference device.

The Quantum Interference Device (QUIT) uses the Aharanov-Bohm Effect but in a slightly different manner than the ABED. The diagram at right shows two quantum island regions "a" and "b" defined by the surrounding electrodes. A magnetic field is applied vertically to the plane of the device. The magnetic field, according to classical physics, tends to force the resident electrons into circular orbits which inhibits the electrons from following a trajectory from region "a" to region "b". However, the magnetic field can be adjusted and the AB effect can be used to cancel out the circular motions and thereby enhance a trajectory passing between the two regions. These were developed in 1994 as a laboratory device.

Josephson Junction

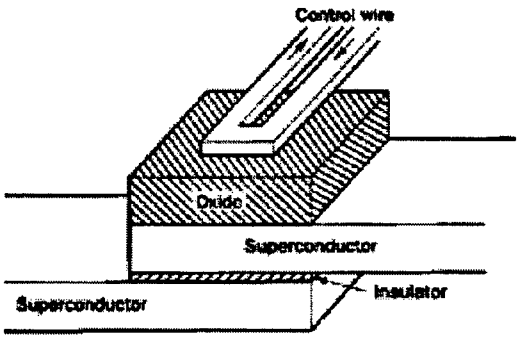


Figure 1.6.5: The Josephson Junction. Picture taken from Turton's book.

A Josephson Junction (JJ) is a superconducting device that uses a weak magnetic field to control the flow of current through a tunneling barrier. A thin oxide layer that serves as the tunneling barrier separates two slabs of superconducting material. A slight voltage is applied across the two slabs so that current will flow through this junction. A small loop of wire functions as a control (or gate) electrode. A small current passing through the loop will produce a small magnetic field which is sufficient to destroy the pairing of charged particles (i.e. suspends superconduction) near the junction. As a result, the resistance of the junction increases and the devices enters the "off" state. JJ devices use up to a factor of 1000 less power than conventional semiconductor devices.

Section 1.8: Discussion of References for this Report

The material for the report has been compiled from a number of books and journals with the author's unique interpretation in some cases. Where possible, the body of the text lists the author or book that can be identified with a specific comment, point of view or derivation. The reference list is divided into two parts. The first lists the books and publications most responsible for the points of view taken in this report. The second provides a sampler of the material already in the literature.

Chapter 2: Review of Elementary Quantum Mechanics and Notation

The application and understanding of quantum electromagnetic states requires familiarity of basic quantum mechanics. This chapter briefly discusses the typical concepts of the classical Hamiltonian, Hilbert space, the quantum mechanical Hamiltonian, representations and the density operator.

Section 2.1: Euclidean Vector Spaces and Dirac Notation

Linear algebra is the natural language of quantum theory. This section reviews the Dirac notation that unifies discrete (i.e., discrete basis sets) and continuous vector spaces with an inner product. First the notation is reviewed for the vector space spanned by the basis set of unit vectors $\{\hat{x},\hat{y},\hat{z}\}$. It is then generalized to function space. The concepts of closure and completeness are discussed.

Topic 2.1.1: Introduction to Inner Product

A (generalized) Hilbert Space H is a vector space with an inner product defined on it. The inner product between two elements f_1 and f_2 in H is denoted by $\langle f_1|f_2\rangle$. The space H is sometimes called an inner product space. Other books use the term Hilbert Space to refer to a vector space of functions that are square integrable in the sense that the following integral exists for all functions $f \varepsilon H$

$$\int_a^b dx\,|f(x)|^2$$

In this books, generalized Hilbert space includes ordinary 3-D vectors and functions.

The inner product (i.e. scalar product) $\langle\bullet|\bullet\rangle$ is closely related to the metric and norm. The Norm is the length of a vector. The metric $d(\vec{r}_1,\vec{r}_2)$ is a "distance function" that measures the distance between two elements $\vec{r}_1,\vec{r}_2$ of the vector space. The norm of a vector is $\|\vec{r}_1\| = \langle\vec{r}_1|\vec{r}_1\rangle^{1/2} = d(\vec{r}_1,0)$.

An inner product $\langle\bullet|\bullet\rangle$ in a (real or complex) vector space F is a scalar valued function that maps $F\times F\rightarrow C$ (where C is the set of complex numbers) with the properties

1. $\langle f|g\rangle = \langle g|f\rangle^*$ with $f,g \varepsilon F$ and "*" denotes complex conjugate
2. $\langle \alpha f+\beta g|h\rangle = \alpha^*\langle f|h\rangle+\beta^*\langle g|h\rangle$ and $\langle h|\alpha f+\beta g\rangle = \alpha\langle h|f\rangle+\beta\langle h|g\rangle$ where $f,g,h \varepsilon F$ and $\alpha,\beta \varepsilon F$
3. $\langle f|f\rangle \geq 0\ \forall f$ and $\langle f|f\rangle = 0$ iff $f=0$ (except at possibly a few points for $C_p[a,b]$)

For function spaces, the inner product and norm are

$$\langle f|g\rangle = \int_a^b dx\, f(x)^* g(x)$$

$$\|f(x)\| = \langle f|f\rangle^{1/2} = \int_a^b dx\, f(x)^* f(x) = \int_a^b dx\, |f(x)|^2$$

As for real vectors, the complex conjugate is also omitted for real functions. The inner product for a function space requires a Rieman integral over the domain of definition. Property 3 holds exactly for continuous functions on [a,b]; however, for piecewise continuous the property is correct except at possibly the points of discontinuity as shown at left. The exceptions arise for piecewise continuous functions because the Rieman integral is insensitive to individual points.

Topic 2.1.2: Kets, Bras and Brackets

The basis vectors for 3D Euclidean space are $\{\hat{x},\hat{y},\hat{z}\}$ which can be written as $\{\hat{e}_1,\hat{e}_2,\hat{e}_3\}$ or $\{\hat{1},\hat{2},\hat{3}$. A "ket", denoted by $|\ \rangle$, is another notation for a vector. The vector $\vec{v}$ can be written as $|v\rangle$ and the basis vectors as

$$\hat{x} \leftrightarrow |1\rangle \quad \hat{y} \leftrightarrow |2\rangle \quad \hat{z} \leftrightarrow |3\rangle$$

A general basis vector appears as

$$\hat{e}_n \leftrightarrow |n\rangle$$

For example, the vector $\vec{v} = 3\hat{x} - 4\hat{y} + 10\hat{z}$ is obviously written as $|v\rangle = 3|1\rangle - 4|2\rangle + 10|3\rangle$. The "ket" is not a component like the "3" or "-4" in the example for $|v\rangle$. By definition, a vector space V contains the sum of two of its vectors $|v\rangle = |v_1\rangle + |v_2\rangle$. Some people write $|v\rangle = |v_1 + v_2\rangle$ which means the ket (or vector) corresponding to the "summed" vector (i.e., $|v_1 + v_2\rangle = |v_1\rangle + |v_2\rangle$). This is somewhat reminiscent of writing, for functions, $(f_1+f_2)(x)=f_1(x)+f_2(x)$. Similarly, $|\alpha v\rangle = \alpha|v\rangle$ where α is a complex constant.

A "bra" $\langle\ |$ is defined to be a projection operator. The bras $\langle 1|, \langle 2|, \langle 3|$ are operators that project a vector $\vec{v}$ onto the unit vectors $\hat{x}$, $\hat{y}$, $\hat{z}$, respectively. For example, if $|v\rangle = 3|1\rangle - 4|2\rangle + 10|3\rangle$ then the projection operations provide the components $\langle 1|\vec{v} = 3$, $\langle 2|\vec{v} = -4$, and $\langle 3|\vec{v} = 10$. The better notation for the combination of projection operators and vectors is

$$\langle 1|\vec{v} = \langle 1|v\rangle$$

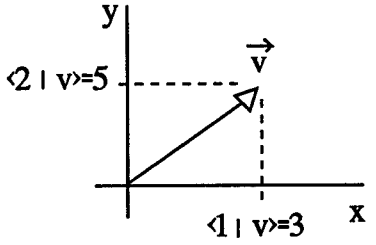


Figure 2.1.1: Projection of $\vec{v} = 3\hat{x} + 5\hat{y}$ onto $|1\rangle, |2\rangle$.

This combination is a "bra" + "ket" which is a "braket" (or bracket). In general, $\langle w|$ is the operator that projects one vector onto the vector $\vec{w}$. Elementary vector analysis says that $\langle w|$ corresponds to $\vec{w}\bullet$ where the dot refers to the usual dot-product

$$\langle w|v\rangle = (\vec{w}\bullet)\vec{v} = \vec{w}\bullet\vec{v}$$

The bras are linear operators and can be distributed across a sum.

$$\langle w|\left[|v_1\rangle+|v_2\rangle\right]=\langle w|v_1\rangle+\langle w|v_2\rangle$$

As a note, other books call the bras "projectors" and they call objects like $|\bullet\rangle\langle\bullet|$ projection operators.

For two general unit vectors, the orthonormality relations are

$$\langle m|n\rangle = \delta_{m,n} = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases} \qquad (2.1.1)$$

Where the Kronecker delta function $\delta_{m,n}$ is defined in Equation 2.1.1. The Kronecker delta function $\delta_{m,n}$ is used to express orthonormality when the basis set is in one-to-one correspondence with a subset of integers (could be an infinite subset). For cases where the set of basis function is in one-to-one correspondence with a dense subset of the real numbers, the Kronecker delta function is replaced by the Dirac Delta Function (i.e., the impulse function) $\delta(x-x')$. A set of vectors $B = \{|1\rangle, |2\rangle, \cdots\}$ is orthonormal if for any two vectors $|m\rangle, |n\rangle$ in B the inner product between them is $\langle m|n\rangle = \delta_{m,n}$.

A linear combination of "n" orthonormal vectors $B = \{|1\rangle, |2\rangle, \cdots, |n\rangle\}$ has the form

$$|v\rangle = \sum_{i=1}^{n} C_i|i\rangle \qquad (2.1.2)$$

where C_i can be complex numbers. The collection of all such vectors $V = \{|v\rangle\}$ forms a vector space and the set B is a basis set. The set B "spans" the vector space V. The dimension of the vector space is Dim(V)=n. Since every vector in V can be found by a suitable choice of the C_i, the set B is said to be complete.

The components of the vector C_i in Equation 2.1.1, can be written in terms of the brackets by projecting the vector $|v\rangle$ onto each basis vector $|m\rangle$.

$$\langle m|v\rangle = \langle m|\sum_{i=1}^{n} C_i|i\rangle = \sum_{i=1}^{n} C_i\langle m|i\rangle = \sum_{i=1}^{n} C_i\delta_{i,m} = C_m \qquad (2.1.3)$$

The results from Equation 2.1.3 written as $C_i = \langle i|v\rangle$, can be substituted into Equation 2.1.2 to obtain

$$|v\rangle = \sum_{i=1}^{n} C_i|i\rangle = \sum_{i=1}^{n}\left[\langle i|v\rangle\right]|i\rangle$$

or

$$|v\rangle = \sum_{i=1}^{n} |i\rangle\langle i|v\rangle \qquad (2.1.4)$$

The *closure or completeness relation*, which obtains from Equation 2.1.4, is of vital importance as a compact expression of a complete set of orthonormal vectors (i.e., a basis set). Equation 2.1.4 can be regrouped

$$|v\rangle = \left(\sum_{i=1}^{n} |i\rangle\langle i|\right)|v\rangle$$

Consider the quantity in parenthesis to be an operator and realize that the equation must hold for all vectors in the vector space V (sort of thinking that the ket $|v\rangle$ can be cancelled on both sides). The last equation becomes

$$\sum_{i=1}^{n} |i\rangle\langle i| = 1$$

for the vector space V spanned by the basis $B = \{|1\rangle, |2\rangle, \cdots, |n\rangle\}$. The "1" that appears in this equation is actually an operator and not just the number "1" . The number of terms in the summation changes for each vector space depending on the number of basis vectors.

Example: The completeness relation for $\mathbb{R}^3$ using $\langle w| = \vec{w}\bullet$ is

$$1 = |1\rangle\langle 1| + |2\rangle\langle 2| + |3\rangle\langle 3|$$

so

$$1 = \hat{x}\hat{x}\bullet + \hat{y}\hat{y}\bullet + \hat{z}\hat{z}\bullet$$

Note that the unit vectors are just written next to each other without an operator between them.

Example: To see that "$\hat{x}\hat{x}\bullet + \hat{y}\hat{y}\bullet + \hat{z}\hat{z}\bullet$" is a unit operator, let $\vec{R} = x\hat{x} + y\hat{y} + z\hat{z}$ and write

$$\begin{aligned}(\hat{x}\hat{x}\bullet + \hat{y}\hat{y}\bullet + \hat{z}\hat{z}\bullet)\vec{R} &= (\hat{x}\hat{x}\bullet + \hat{y}\hat{y}\bullet + \hat{z}\hat{z}\bullet)(x\hat{x} + y\hat{y} + z\hat{z}) \\ &= (\hat{x}\hat{x}\bullet + \hat{y}\hat{y}\bullet + \hat{z}\hat{z}\bullet)x\hat{x} + (\hat{x}\hat{x}\bullet + \hat{y}\hat{y}\bullet + \hat{z}\hat{z}\bullet)y\hat{y} + (\hat{x}\hat{x}\bullet + \hat{y}\hat{y}\bullet + \hat{z}\hat{z}\bullet)z\hat{z} \\ &= x\hat{x} + y\hat{y} + z\hat{z}\end{aligned}$$

So we get the vector $\vec{R}$ back again. The operator "$\hat{x}\hat{x}\bullet + \hat{y}\hat{y}\bullet + \hat{z}\hat{z}\bullet$" just decomposes $\vec{R}$ into components and then re-assembles it again.

Topic 2.1.3: The Euclidean Dual Vector Space

The previous subtopic shows that a bra $\langle w|$, which projects an arbitrary vector onto the vector $\vec{w}$, is an operator that maps a vector space V into the complex numbers C (i.e., $\langle w| : V \rightarrow C$). The bra $\langle w|$ is a linear operator and can be distributed across terms. These projection operators form a vector space – a vector space of linear operators.

The space "dual" to the vector space V is defined to be the set V^+ consisting of all of the bra operators $\langle w|$. For each ket $|w\rangle$, there is a bra $\langle w|$ and vice versa. Each element of the original vector space V is in 1-1 correspondence with the "dual vector space V^+". Mathematically, the two vector spaces $V = \{|v\rangle\}$ and $V^+ = \{\langle w|\}$ are related by an anti-linear 1-1 (isomorphic) map denoted by the dagger superscript. The isomorphic map $+ : V \leftrightarrow V^+$ is call the Hermitian conjugate (or adjoint operator).

$$\langle \bullet| \underset{+}{\leftrightarrow} |\bullet\rangle$$

The usual way to write this mapping is as

$$|w\rangle^+ = \langle w|$$

If $\alpha, \beta \in C$ (the complex numbers) then the anti-linearity property is written as

$$\left[\alpha|v\rangle + \beta|w\rangle\right]^+ = \left[\alpha|v\rangle\right]^+ + \left[\beta|w\rangle\right]^+ = \alpha^*\langle v| + \beta^*\langle w|$$

where " * " indicates complex conjugate. Part of the reason for taking the complex conjugate of the coefficients has to do with finding the magnitude of a "complex" vector.

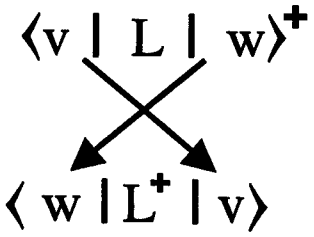
The adjoint operator maps a basis set for V into a corresponding basis set for V^+. If $\{|i\rangle : i = 1,\ldots,n\}$ is a basis set for V then $\{\langle i| : i = 1,\ldots,n\}$ is a basis set for V^+. Therefore the dual basis set is composed of operators that project an arbitrary vector onto the set of basis vectors of the vector space V.

Example: Some relations can be demonstrated for $\vec{v} = |v\rangle = a|1\rangle + b|2\rangle$ where $\{|1\rangle,|2\rangle\}$ spans R^2.

1. $\langle v| = |v\rangle^+ = \left[a|1\rangle + b|2\rangle\right]^+ = a^*\langle 1| + b^*\langle 2|$
2. $\langle v|1\rangle = \left[a^*\langle 1| + b^*\langle 2|\right]|1\rangle = a^*$ and $\langle 1|v\rangle = \langle 1|\left[a|1\rangle + b|2\rangle\right] = a$
3. $\langle 1|v\rangle = a = (a^*)^* = \langle v|1\rangle^*$. Note that $\langle v|1\rangle^+ = \langle 1|v\rangle = \langle v|1\rangle^*$.

For the sake of argument, let L, L_1, L_2 be linear operators that act on the vector space V with basis vectors $\{|1\rangle,|2\rangle,\ldots,|n\rangle\}$ where $L=L_1L_2$ (for example). Note that L does not need to be a projection operator (there are other types of operators); however, as seen in a subsequent section the linear operators can be written as combinations of bras and kets. As a prototype consider $\langle w|L|v\rangle$ where $|v\rangle,|w\rangle \in V$ and $\langle w| \in V^*$. The adjoint operator reverses the direction of all the objects and adds the "+" to each operator.

$$\langle v|L_1L_2|w\rangle^+ = \langle w|L_2^+L_1^+|v\rangle.$$



Topic 2.1.2: Action of the adjoint operator.

Topic 2.1.4: Inner Product and Norm

Assume $\{|i\rangle : i = 1,2,3$ is a basis set for a 3D vector space. The norm (or length) of a vector is found by taking the square root of the inner product.

$$\|\vec{v}\|^2 = \vec{v} \bullet \vec{v} = \langle v|v\rangle$$

$$= \left(\sum_{i=1}^{3} v_i |i\rangle\right)^{+} \left(\sum_{j=1}^{3} v_j |j\rangle\right) = \sum_{i=1}^{3} \langle i| v_i^* \sum_{j=1}^{3} v_j |j\rangle$$

$$= \sum_{i,j=1}^{3} \langle i| v_i^* v_j |j\rangle = \sum_{i,j=1}^{3} v_i^* v_j \langle i|j\rangle$$

The last step follows since $v_i^* v_j$ is just a number and so it can be moved outside the brackets. Now we use the orthonormality property for unit vectors to write

$$\|v\|^2 = \sum_{i,j=1}^{3} v_i^* v_j \delta_{i,j} = \sum_{i=1}^{3} v_i^* v_i = \sum_{i=1}^{3} |v_i|^2$$

where $|v_i|$ is the magnitude of the complex number.

Section 2.2: Hilbert Space of Functions

The first topic shows how the brackets allow functions to be thought of as components of a vector. The next sections shows the Hilbert Space of functions where a function can be written as a linear combination of basis functions. Although the two representations are related, one thinks of projecting a function into coordinate space as different from a projection into a Hilbert space with basis functions. These two sections will compare and contrast the concepts.

We will provide examples using Fourier series. Sections at the end of this chapter will examine some issues of convergence.

The inner product is formed by the combination of a "bra" and a "ket". The inner product is thought of in two different ways for functions even though they are related. This section reviews vector spaces, basis sets and closure relations for functions. The inner product $\langle x|f\rangle$ where "x" refers to the x-axis is really the function f(x). It follows from the definition of inner product when "x" is taken to mean the Dirac delta function (impulse function). This will be seen in more detail.

Topic 2.2.1: Vector Space of Functions

Functions in a set $F=\{\phi_o,\phi_1,\phi_2,\ldots,\phi_n\}$ are linearly independent if for complex constants c_i (i=0,...,n}, the sum

$$\sum_{i=0}^{n} c_i\phi_i(x)=0$$

can only be true when all of the complex constants are zero c_i=0. Functions in the set $F=\{\phi_o,\phi_1,\phi_2,\ldots,\phi_n\}$ are orthonormal if $\langle\phi_i|\phi_j\rangle=\delta_{ij}$ for every integer i,j in the set {0,1,2,...,n}. The inner product between two complex functions "f" and "g" over the range (a,b) is

$$\langle f|g\rangle=\int_a^b dx\, f^*(x)g(x)$$

Orthonormal functions are linearly independent as can be seen by starting with a sum over the functions ϕ_i using complex coefficients c_i

$$\sum c_i\phi_i=0 \quad \text{or} \quad \sum c_i|\phi_i\rangle=0$$

Next, operate on both sides with the bra operator $\langle\phi_m|$ to get

$$\sum c_i\langle\phi_m|\phi_i\rangle=0$$

but

$$\delta_{mi}=\langle\phi_m|\phi_i\rangle$$

so that the complex coefficients must all be 0 (i.e., c_i=0 for all i).

A linearly independent set of functions $F=\{\phi_o,\phi_1,\phi_2,\ldots,\phi_n\}$ is complete if every function f(x) in the space can be written as

$$f(x)=\sum_{i=1}^{\infty}c_i\phi_i(x) \quad \text{or} \quad |f\rangle=\sum_{i=1}^{\infty}c_i|\phi_i\rangle$$

(except at possibly a few points) for some choice of complex numbers c_i. If the set $\{\phi_i\}$ is "complete and orthonormal" then the functions ϕ_i are basis functions (or basis vectors) and span the function space. The complete orthonormal set of functions $F=\{\phi_o,\phi_1,\phi_2,\ldots,\phi_n\}$ form of basis for a Hilbert space H. The basis functions can be written using Dirac notation as $\{|\phi_o\rangle,|\phi_1\rangle,\ldots\}$ or, more conveniently as $\{|0\rangle,|1\rangle,\ldots\}$. Assume that the infinite series uniformly converge so that they can be integrated or differentiated as necessary.

Results very similar to that for the Euclidean space (actually, Euclidean space is a misnomer since function-space also has orthonormal basis vectors!) can be written for the function the Hilbert space H. The basis vector expansion is

$$|f\rangle=\sum_{i=0}^{\infty}c_i|i\rangle$$

The components of the vector $|f\rangle$ (i.e., the expansion coefficients c_i) can be found by operating with the bra $\langle j|$ as follows

$$\langle j|f\rangle=\langle\phi_j|f\rangle=\langle j|\sum_{i=0}^{\infty}c_i|i\rangle=\sum_{i=0}^{\infty}c_i\langle j|i\rangle=\sum_{i=0}^{\infty}c_i\delta_{ij}=c_j$$

so, just like Euclidean vectors, we find the vector components to be $c_j=\langle j|f\rangle$. The definition of the vector components $c_i=\langle\phi_i|f\rangle=\langle i|f\rangle$ can be used to demonstrate a closure relation.

$$|f\rangle=\sum_{i=0}^{\infty}c_i|i\rangle=\sum_{i=0}^{\infty}\langle i|f\rangle|i\rangle=\sum_{i=0}^{\infty}|i\rangle\langle i|f\rangle=\left(\sum_{i=0}^{\infty}|i\rangle\langle i|\right)|f\rangle$$

The closure relation is found by comparing both sides

$$\sum_{i=0}^{\infty}|i\rangle\langle i|=1$$

The closure relation is equivalent to saying that the set $F=\{\phi_o,\phi_1,\phi_2,\ldots,\phi_n\}$ is complete.

The bra for functions can be written in terms of an operator as

$$\langle f|\equiv\int_a^b dx\,f^*\circ$$

where the small circle is a place holder for another function.

Topic 2.2.2: Projecting Functions into Coordinate Space

Recall that Euclidean vector $\vec{v}$ (Section 2.5) in a generalized Hilbert space has components v_i. The components are really functions of the index "i" as in $v(i) = v_i = \langle i|v\rangle$. This is equivalent to projecting the vector $\vec{v}$ on to the i^{th} coordinate. The index "i" is thought of similarly to the x-axis, for example, except that "i" refers to the integer subset of the reals.

The function "f" in a vector space is written as $|f\rangle$. The function "f" (i.e., $|f\rangle$) is regarded as a physical object and not the component f(x). The reason is that "f" can be represented, for example, as f(x), or as the Fourier transform f(k) or as a series expansion. All three representations are equally valid expression for "f". The component f(x) is equivalent to projecting the function "f" onto the x^{th} coordinate. The Fourier transform is equivalent to projecting "f" into k-space to get $\langle k|f\rangle = f(k)$. The same "f" is used for f(k) as for f(x) with the understanding that the explicit form of the projected function is different (i.e., f(k) is not found by replacing "x" with "k").

The connection between the definition of $\langle g|h\rangle$ as an integral and the interpretation of $\langle x|f\rangle$ as the function f(x) is obtained through the Dirac delta function. If x_o is a coordinate, then ket $|x_o\rangle$ is actually the delta function $\delta(x-x_o)$. Then the inner product (i.e., the projection of "f" onto x_o) is

$$\langle x_o|f\rangle = \langle \delta(x-x_o)|f(x)\rangle = \int_{-\infty}^{\infty} dx\,[\delta(x-x_o)]^* f(x) = \int_{-\infty}^{\infty} dx\ \delta(x-x_o) f(x) = f(x_o)$$

The vector $\vec{f} = |f\rangle$ is therefore decomposed into the delta-function basis vectors $\{\delta(x-x_o) : x_o \in \mathbb{R}\}$ of which there are an infinite number.

The closure relation for Euclidean vectors can be generalized for use with coordinate space. The summation used with discrete basis vectors $|i\rangle$ is replaced with an integral for the continuous set of basis vectors such as $|x\rangle$. The basis set for the Fourier transform set is also continuous and the summation is also replaced with an integral. The relations for closure and orthonormality become

$$\sum_{i=1}^{n} |i\rangle\langle i| = 1 \quad \rightarrow \quad \int |x\rangle dx \langle x| = 1$$

$$\langle m|n\rangle = \delta_{mn} \quad \rightarrow \quad \langle x'|x\rangle = \delta(x-x')$$

$$m, n \in \text{integers} \qquad x, x' \in \mathbb{R}$$

The relations can be demonstrated using the delta-function definition of the coordinate ket $|x\rangle$. The proofs are left to the reader. Notice in particular the new representation for the Dirac delta function.

To see that the notation is consistent, operate on the left of

$$1 = \int |x\rangle dx \langle x|.$$

with the bra $\langle x'|$ and on the right by a function $|f\rangle$ to get

$$\langle x'|f\rangle = \langle x'|f\rangle = \langle x'|1|f\rangle = \langle x'|\int |x\rangle dx \langle x||f\rangle$$
$$= \int \langle x'|x\rangle dx \langle x|f\rangle = \int \delta(x-x')f(x)\,dx$$
$$= f(x')$$

Generally, closure relations are used to write inner products in a more tractable form. For example, considered two functions $|f\rangle, |g\rangle$. The inner product is

$$\langle g|f\rangle = \langle g|1|f\rangle = \langle g|\int |x\rangle dx \langle x||f\rangle = \int \langle g|x\rangle dx \langle x|f\rangle$$

Now use $\langle x|g\rangle^+ = \langle g|x\rangle$ and the fact that $\langle x|g\rangle$ is a complex number so that $\langle x|g\rangle^+ = \langle x|g\rangle^*$ and then

$$\langle g|f\rangle = \int dx\, g^*(x) f(x)$$

as required by the basic definition of inner product.

The Dirac delta function is used when the basis vectors, such as $|x\rangle$, are continuous functions. When the index "i" (as in $|i\rangle$) is discrete, the Kronecker delta function is used in place of the Dirac delta function. The index "i" does not need to be an integer so long as it's range of values is in one-to-one correspondence with a subset of the integers. The use of the Dirac or Kronecker delta function is the only difference between using continuous or discrete basis functions, respectively.

Topic 2.2.3: Miscellaneous Relations

The norm (squared) of a function can be calculated

$$\|f\|^2 = \langle f|f\rangle = \langle f|1|f\rangle = \langle f|\left[\int |x\rangle dx \langle x|\right]|f\rangle = \int \langle f|x\rangle dx \langle x|f\rangle = \int \langle x|f\rangle^+ dx \langle x|f\rangle$$
$$= \int f^*(x) f(x) dx = \int |f(x)|^2$$

Each different set of basis vectors F leads to a different representation of the Dirac delta function. For example, using $\delta(x-x') = \langle x|x'\rangle$

$$\delta(x-x') = \langle x|x'\rangle = \langle x|1|x'\rangle$$
$$= \langle x|\left[\sum_{i=0}^{\infty} |\phi_i\rangle\langle\phi_i|\right]|x'\rangle$$
$$= \sum_{i=0}^{\infty} \langle x|\phi_i\rangle\langle\phi_i|x'\rangle$$
$$= \sum_{i=0}^{\infty} \phi_i^*(x)\phi_j(x')$$

Section 2.3: Linear Operators

This chapter discusses linear transformations on a Hilbert space. These linear operators have matrix representations. Unitary operators change the basis. The adjoint operator is discussed along with the self adjoint and Hermitian operators. This section denotes operators by a caret above the symbol as in $\hat{O}$ and the corresponding matrices with an underline $\underline{O}$. Linear operators can be represented by matrices once a basis set is identified for the vector space.

Topic 2.3.1 Linear Transformations and Matrices

A linear transformation T defined on a Hilbert space maps one vector $|v_a\rangle$ in the space into another $|v_c\rangle$ according to

$$T|v_a\rangle = |v_c\rangle$$

If $|v_a\rangle$ and $|v_b\rangle$ are elements of the Hilbert space and c_1, c_2 are complex numbers (C) then

$$T[c_1|v_a\rangle + c_2|v_b\rangle] = c_1 T|v_a\rangle + c_2 T|v_b\rangle$$

where the new vectors $T|v_a\rangle$ and $T|v_b\rangle$ are also members of the Hilbert space.

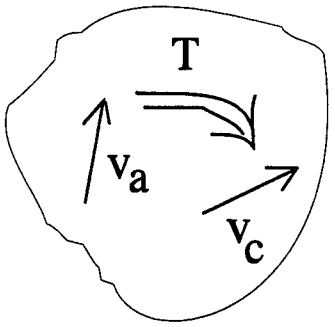


Figure 2.3.1: Linear operator maps vectors.

A linear transformation T can be represented by a matrix. The transformation "T" (i.e., linear operator) maps elements of the Hilbert space into other elements of the Hilbert space. However, each element of the Hilbert space is a vector which is a linear combination of the "unit vectors" in the basis set. Knowing how "T" maps each basis vector is sufficient to know how "T" maps all vectors $|v\rangle$.

For simplicity, consider a linear transformation "T" that maps a vector space into itself $T: V \to V$ so that the matrix is square. Let V be an n-dimensional (generalized) Hilbert space with basis $B = \{\phi_i = |\phi_i\rangle = |i\rangle : i = 1,2,\ldots,N\}$. The matrix of the operator "T" with respect to the basis set B is

$$\underline{T}_{ij} = \begin{bmatrix} T_{11} & T_{12} & \cdots & T_{1N} \\ T_{21} & T_{22} & \cdots & T_{2N} \\ \vdots & & & \vdots \\ T_{N1} & T_{N1} & \cdots & T_{NN} \end{bmatrix} = \underline{T}$$

where T_{ij} occurs in

$$T\phi_j = \sum_i T_{ij}\phi_i \qquad (2.3.1)$$

Note the order of i,j on the matrix element T_{ij}. Equation 2.3.1 can also be written as

$$T|\phi_j\rangle = \sum_i T_{ij}|\phi_i\rangle \qquad \text{or} \qquad T|j\rangle = \sum_i T_{ij}|i\rangle$$

The non-square matrix is similar except two different vector spaces are required.

The matrix elements are most conveniently written in Dirac notation. For simplicity, consider an operator that maps a vector space into itself according to $T: V \to V$. As before, assume V is spanned by the basis vectors (Euclidean or functions) $B_v = \{|\phi_i\rangle = |i\rangle : i = 1,2,\ldots,N\}$

$$T|\phi_a\rangle = \sum_i T_{ia}|\phi_i\rangle$$

Operating with a projection operator $\langle\phi_b|$, we have

$$\langle\phi_b|T|\phi_a\rangle = \langle\phi_b|\sum_i T_{ia}|\phi_a\rangle = \sum_i T_{ia}\langle\phi_b|\phi_i\rangle = \sum_i T_{ia}\delta_{bi} = T_{ba}$$

So inner products involving unit vectors and a linear transformation T are really elements of a matrix. Note the order of the indices b,a.

As a note, the adjoint of a matrix is

$$\underline{M}^+ = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^+ = \begin{bmatrix} a^* & c^* \\ b^* & d^* \end{bmatrix}$$

Topic 2.3.2: Basis Vector Representation of a Linear Operator

It is sometimes useful to represent an operator T in terms of bras and kets (i.e., in terms of the basis vectors for the vector space and the dual space). First consider an operator $T: V \to V$ that maps of vector space V into itself with basis $B_v = \{|i\rangle\}$. The definition of matrix gives

$$T|b\rangle = \sum_a T_{ab}|a\rangle$$

where $|a\rangle$ and $|b\rangle$ are basis vectors. Multiplying this last equation by $\langle b|$ *from the right,* provides

$$T|b\rangle\langle b| = \sum_a T_{ab}|a\rangle\langle b|$$

Now sum both sides over the index "b"

$$T\sum_b |b\rangle\langle b| = \sum_{a,b} T_{ab}|a\rangle\langle b|$$

The closure relation $\sum|b\rangle\langle b| = 1$ provides

$$T = \sum_{a,b} T_{ab}|a\rangle\langle b|$$

These basis vector representations of an operator have a form very reminiscent of the closure relation. In fact, the closure relation obtains if the operator T is taken as the unit operator T=1 so that the matrix elements are $T_{ab}=\delta_{ab}$.

Example: For the linear operator $T: V \to V$ find an operator that maps the basis vectors as follows

$$|1\rangle \to |2\rangle \qquad \text{and} \qquad |2\rangle \to -|1\rangle$$

The solution can be found by noting that $|2\rangle\langle 1|$ can operate on the unit vector $|1\rangle$ and it gives $|2\rangle\langle 1|1\rangle = |2\rangle$. We similarly notice that $\left(-|1\rangle\langle 2|\right)|2\rangle = -|1\rangle\langle 2|2\rangle = -1$. We therefore speculate that the desired operator is

$$T = |2\rangle\langle 1| - |1\rangle\langle 2|$$

The reader can try the operator on both basis vectors. The first basis vector

$$T|1\rangle = \left(|2\rangle\langle 1| - |1\rangle\langle 2|\right)|1\rangle = |2\rangle$$

The transformation T describes her rotation by 90°.

Topic 2.3.3: Direct Product Space

Vector spaces V and W can be combined into product spaces with basis vectors of the form

$$\left\{|\phi_i\rangle|\Psi_j\rangle = |\phi_i, \Psi_j\rangle\right\}$$

where the individual spaces have the basis vectors

$$B_v = \left\{|\phi_i\rangle\right\} \qquad B_w = \left\{|\Psi_j\rangle\right\}$$

and the spaces V and W do not need to be the same size. The size of the direct product space $V \otimes W$ is given by

$$\text{Dim}[V \otimes W] = \text{Dim}(V)\,\text{Dim}(W).$$

The projection operators (bras) corresponding to

$$|v,w\rangle \in V \otimes W \text{ are } |v,w\rangle^+ = \langle v,w| = \langle v|\langle w|$$

where $\langle v,w| \in [V \otimes W]^*$. The basis set for the dual space is

$$\left\{\langle\phi_i, \Psi_j| = \langle\phi_i|\langle\Psi_j|\right\}$$

Next, consider inner products on the product space. Inner products can only be formed between V* and V, and also between W* and W. So if

$$|v_1\rangle, |v_2\rangle \in V \qquad |w_1\rangle, |w_2\rangle \in W$$

then the inner product is

$$\langle v_1 w_1 | v_2 w_2\rangle = \langle v_1|v_2\rangle\langle w_1|w_2\rangle \qquad (2.3.2)$$

Of course, $\langle v_1|v_2\rangle$ and $\langle w_1|w_2\rangle$ are just complex numbers, and (2.3.2) can also be written as

$$\langle v_1 w_1 | v_2 w_2\rangle = \langle w_1|w_2\rangle\langle v_1|v_2\rangle$$

where the factors on the right hand side have been reversed.

There are other operators O that act on the direct product space. The matrix of O can be defined by

$$O|\phi_a \Psi_b\rangle = \sum_{\alpha,\beta} O_{\alpha,\beta;a,b}|\phi_\alpha \Psi_\beta\rangle$$

or, taking the inner product using the projection operator $\langle \phi_c, \Psi_d |$, provides

$$\langle \phi_c \Psi_d | O | \phi_a \Psi_b \rangle = \sum_{\alpha,\beta} O_{\alpha,\beta;a,b} \langle \phi_c \Psi_d | \phi_\alpha \Psi_\beta \rangle = \sum_{\alpha,\beta} O_{\alpha,\beta;a,b} \delta_{c\alpha} \delta_{d\beta} = O_{cd;ab}$$

Topic 2.3.4: Trace of an Operator

The trace of an operator is essentially defined as the trace of its matrix, which is the sum of the diagonal elements. Let $\{|a\rangle\}$ be the set of basis vectors for the vector space V. Let $T : V \to V$ be a linear operator. The trace of T is

$$\mathrm{Tr}(T) = \sum_a \langle a | T | a \rangle$$

Recall that the trace is independent of the particular choice of basis vectors

The following lists some useful properties of the trace. Assume that the operators A, B, C have a domain and range within a single vector space V with basis vectors $B_v = \{|a\rangle\}$

(1) Tr(AB)=Tr(BA)

proof:

$$\mathrm{Tr}(AB) = \sum_n \langle n | AB | n \rangle = \sum_n \langle n | A1B | n \rangle = \sum_{nm} \langle n | A | m \rangle \langle m | B | n \rangle = \sum_{nm} \langle m | B | n \rangle \langle n | A | m \rangle$$

$$= \sum_m \langle m | BA | m \rangle = TR(BA)$$

(b) TR(ABC) = TR(BCA) = TR(CAB)

Topic 2.3.5: Trace in Coordinate Space

Starting with the definition of trace, inserting a unit operator in two places, and then the closure relation in coordinate space gives

$$\mathrm{Tr}\,\hat{A} = \sum_n \langle n | \hat{A} | n \rangle = \sum_n \langle n | 1\hat{A}1 | n \rangle = \iint dx\,dx' \sum_n \langle n | x \rangle \langle x | \hat{A} | x' \rangle \langle x' | n \rangle$$

The matrix elements are number that can be rearranged to give

$$\mathrm{Tr}\,\hat{A} = \iint dx\,dx' \sum_n \langle x | \hat{A} | x' \rangle \langle x' | n \rangle \langle n | x \rangle = \iint dx\,dx' \langle x | \hat{A} | x' \rangle \langle x' | x \rangle = \iint dx\,dx' \langle x | \hat{A} | x' \rangle \delta(x - x')$$

where the closure relation is used for $|n\rangle$ and the Dirac delta function is substituted. Performing the final integration gives

$$\mathrm{Tr}\,\hat{A} = \int dx \langle x | \hat{A} | x \rangle$$

Section 2.4: Unitary Operators and Similarity Transformations

This section discusses operators with a domain and range within a single vector space. Unitary operators preserve angles and lengths. Therefore, if a unitary operator acts on a set of basis vectors, the image set is another set of basis vectors. The unitary operator is critical to understanding the evolution operator obtained from Schrodinger's equation. The orthogonal operators, which rotate Euclidean vectors) are subsets of the unitary operators. The discussion freely interchanges the focus from linear operators to matrices and back owing to the natural isomorphism existing between the space of operators and the space of matrices.

Topic 2.4.1: Orthogonal Rotation Matrices

Orthogonal matrices rotate Euclidean vectors. The word "orthogonal" is meant to imply that the length of a vector does not change when the coordinates are rotated. The orthogonal operator is most conveniently defined through its matrix. An orthogonal matrix $\underline{R}$ has the property that the inverse is related to the transpose as

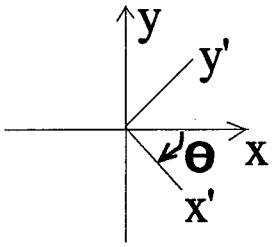


Figure 2.4.1: Rotation of axes

$$\underline{R}^{-1} = \underline{R}^{T} \qquad (2.4.1)$$

This relation is independent of the basis set chosen for the vector space (as it should be since the *operator* does not depend on the chosen basis set). Recall the definition of the transpose

$$\left(R^{T}\right)_{ab} = R_{ba} \quad \text{or} \quad R^{T}_{ab} = R_{ba}$$

where the indices "a" and "b" are interchanged. The defining relation in Equation 2.4.1 can be used to show $\mathrm{Det}\left(\hat{R}\right) = 1$.

Recall that rotation can be viewed as either rotating the coordinate system or as rotating a vector $\vec{r}$. As an example, consider a rotating all vectors by 45° counter clockwise and then re-expressing them in terms of the original basis vectors $|1\rangle$ and $|2\rangle$. The rotation operator provides $\hat{R}|1\rangle = |1'\rangle$. However, the definition of a matrix requires $\hat{R}|1\rangle = R_{11}|1\rangle + R_{21}|2\rangle$. Therefore

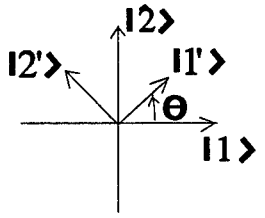


Figure 2.4.2: Rotating the basis vectors and re-expressing them in the original basis set.

$$|1'\rangle = R|1\rangle = \frac{1}{\sqrt{2}}|1\rangle + \frac{1}{\sqrt{2}}|2\rangle = \cos\theta\,|1\rangle + \sin\theta\,|2\rangle = R_{11}\,|1\rangle + R_{21}\,|2\rangle$$

$$|2'\rangle = R|2\rangle = -\frac{1}{\sqrt{2}}|1\rangle + \frac{1}{\sqrt{2}}|2\rangle = -\sin\theta\,|1\rangle + \cos\theta\,|2\rangle = R_{12}\,|1\rangle + R_{22}\,|2\rangle$$

where the coefficients are obtained from the figure. The results can be written as

$$\begin{aligned} R &= R_{11}|1\rangle\langle 1| + R_{12}|1\rangle\langle 2| + R_{21}|2\rangle\langle 1| + R_{22}|2\rangle\langle 2| \\ &= \cos\theta\,|1\rangle\langle 1| + \sin\theta\,|1\rangle\langle 2| - \sin\theta\,|2\rangle\langle 1| + \cos\theta\,|2\rangle\langle 2| \end{aligned}$$

The orthogonal linear operators preserve the orthonormality relation and therefore transform one basis set into another. This can be seen as follows for a 2-D space with basis vectors $|1\rangle,|2\rangle$. Let

$$\hat{R}|a\rangle = |v_a\rangle \qquad \hat{R}|b\rangle = |v_b\rangle$$

where a,b are 1 or 2 for the two basis vectors. Forming the inner product provides

$$\langle v_a|v_b\rangle = \left(|v_a\rangle\right)^+|v_b\rangle = \left(\hat{R}|a\rangle\right)^+\hat{R}|b\rangle = \langle a|\hat{R}^+\hat{R}|b\rangle$$

However, R is real and so $R^+=R^{*T}=R^T$

$$\langle v_a|v_b\rangle = \langle a|\hat{R}^+\hat{R}|b\rangle = \langle a|\hat{R}^T\hat{R}|b\rangle = \langle a|1|b\rangle = \langle a|b\rangle = \delta_{ab}$$

where the defining relation for orthogonal operaters $R^T=R^{-1}$ is used for the forth step. As a result, R preserves the orthonormality relations among the vectors. In fact, orthogonal rotations preserve the lengths of all vectors and all angles -- not just those for the basis vectors.

Topic 2.4.2: Unitary Transformations

A unitary transformation is a "rotation" in the generalized Hilbert space. A unitary operator "u" is defined to have the property that

$$u^+ = u^{-1}$$

This unitary property leads to $|\det(u)|^2 = 1$ by taking the determinant of both sides of

$$uu^+ = 1$$

Assume that the operator u maps the vector space V into itself $u: V \to V$. Unitary operators preserve the orthonormality relations of the basis set. That is, if

$$B_v = \{|a\rangle\}$$

is a basis set then so is

$$B'_v = \{u|a\rangle = |a'\rangle\}$$

The proof is similar to that for orthogonal operators.

$$\langle a'|b'\rangle = \left(u|a\rangle\right)^+\left(u|b\rangle\right) = \langle a|u^+u|b\rangle = \langle a|1|b\rangle = \langle a|b\rangle = \delta_{ab}$$

As a result, B'_v and B_v are equally good basis sets for the Hilbert space V.

The inverse of the unitary operator "u", $u^{-1}=u^+$ can be written in matrix rotation as

$\underline{u}^+ = \underline{u}^{T*}$ or $\left(u^+\right)_{ab} = u^*_{ba}$ or sometimes $u^+_{ab} = u^*_{ba}$

Example: If $u = \sum_{ab} u_{ab}|a\rangle\langle b|$ then u^+ can be calculated as

$$u^+ = \sum_{ab}\left(u_{ab}|a\rangle\langle b|\right)^+ = \sum_{ab}\left(u_{ab}\right)^+|b\rangle\langle a|$$

The coefficient u_{ab} is a single complex number and not the entire matrix so that the dagger can be replaced by the complex conjugate without interchanging the indices. As a result

$$u^+ = \sum_{ab} u^*_{ab} |b\rangle\langle a|$$

The result can be checked to see that $u^+u=1$.

$$u^+u = \left(\sum_{\alpha\beta} u^*_{\alpha\beta} |\beta\rangle\langle\alpha|\right)\left(\sum_{ab} u_{ab} |a\rangle\langle b|\right) = \sum_{\substack{ab \\ \alpha\beta}} u^*_{\alpha\beta} u_{ab} |\beta\rangle\langle b| \delta_{a\alpha} = \sum_{\substack{ab \\ \beta}} u^*_{a\beta} u_{ab} |\beta\rangle\langle b|$$

Working with the product of the unitary matrices provides

$$\sum_a u^*_{a\beta} u_{ab} = \sum_a \left(u^+\right)_{\beta a} u_{ab} = \left(u^+u\right)_{\beta b} = \delta_{\beta b}$$

Notice that the indices are switched for Hermitian adjoint of the matrix since the entire matrix is referenced. Substituting this result for the unitary matrices gives the results

$$u^+u = \sum_{\beta b} \delta_{\beta b} |\beta\rangle\langle b| = \sum_b |b\rangle\langle b| = 1$$

Topic 2.4.3: A Convenient Picture for Unitary Transformations

The unitary linear operator maps replaces one basis set with another.

$$B_v = \{|a\rangle\} \rightarrow B'_v = \{u|a\rangle = |a'\rangle\}$$

For example, the figure shows

$$u|1\rangle = |1'\rangle \qquad u|2\rangle = |2'\rangle$$

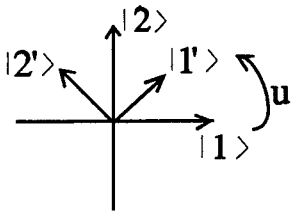


Figure 2.4.3: Rotation by an unitary operator

On the other hand, the two objects $|1'\rangle\langle 1|$ and $|2'\rangle\langle 2|$ perform the following mappings:

$|1'\rangle\langle 1|$ maps $|1\rangle \rightarrow |1'\rangle$ since $\left[|1'\rangle\langle 1|\right]|1\rangle = |1'\rangle\langle 1|1\rangle = |1'\rangle$

$|2'\rangle\langle 2|$ maps $|2\rangle \rightarrow |2'\rangle$ since $\left[|2'\rangle\langle 2|\right]|2\rangle = |2'\rangle\langle 2|2\rangle = |2'\rangle$

Notice that the objects $|1'\rangle\langle 1|$ and $|2'\rangle\langle 2|$ "basis vectors" for the vector space of operators $\{u\}$. Putting both pieces together gives

$$u = |1'\rangle\langle 1| + |2'\rangle\langle 2|$$

It is useful to write "u" as

$$u = \sum_a |a'\rangle\langle a|$$

to handle "rotations" in all dimensions. When using "u" for actual calculations, either $|a'\rangle$ must be expressed as a sum over $|a\rangle$ or vice versa.

Section 2.5: Hermitian Operators and the Eigenvector Equation

The concepts of adjoint, self-adjoint and Hermitian operators are essential for the study of Sturm-Liouville problems in mathematics and quantum mechanics. This book assumes that self-adjoint and Hermitian operators are the same. Hermitian operators have valuable properties that govern the allowable eigenfunctions and eigenvalues. Eigenfunctions are generally used as the basis set for the vector space.

In quantum mechanics, one of the most basic assumptions is that a Hermitian operator can represent every physically observable quantity such as energy, momentum and electric field. An "observable" is the physical object whereas the "Hermitian operator" is the mathematical object.

The result of measuring an observable is assumed to be an eigenvector of the operator. For example, let $\hat{H}$ be the energy operator and for now assume that the Hilbert space is spanned by $\{|n\rangle\}$ the complete set of eigenvectors of $\hat{H}$. Then $\hat{H}|n\rangle = E_n|n\rangle$ says that the observation of the energy of a particle in the state $|n\rangle$ gives the values of the allowed energies E_n. The expected value of the Hamiltonian $\hat{H}$ of a particle occupying the state $|n\rangle$ as

$$\langle n|\hat{H}|n\rangle = \langle n|E_n|n\rangle = E_n$$

Clearly, the notion of Hermitian operator is of immense importance to the physical world especially given that all physically observable quantities correspond to a Hermitian operator. Based on this fact, it seems reasonable that the result of any observation must somehow be contained in the vectors occupying the Hilbert space. In particular, the results must be contained in the basis vectors. Apparently the "completeness" of a basis set is related to a "completeness" in nature.

Topic 2.5.1: Adjoint, Self-Adjoint and Hermitian Operators

Let T be a linear transformation defined on Hilbert space V where $T: V \to V$ (for simplicity) and let the basis vectors be $\{|n\rangle : n = 1,2,\ldots\}$. The adjoint operator T^+ is defined to be the operator which satisfies

$$\langle g|Tf\rangle = \langle T^+ g|f\rangle$$

for all functions $|f\rangle$ and $|g\rangle$ in the Hilbert space.

For functions, we will see that T^+ depends on the boundary conditions that help to define the Hilbert space. In our study so far, we have assumed that T^+ exists and we "understand" it through our understanding of the dual vector space.

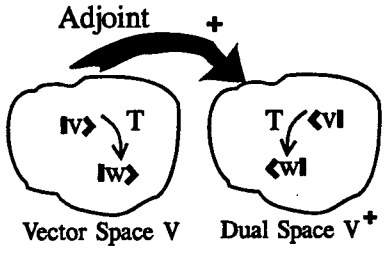


Figure 2.9.1: The adjoint map

Definition: An operator O is Hermitian (i.e., self-adjoint) when $O^+ = O$. The term "Hermitian" is often reserved for complex Hilbert spaces.

Example: If $\hat{O} = \frac{\partial}{\partial x}$ then find O^+ for the Hilbert space of differentiable functions that approach zero as $x \to \pm\infty$. The Hilbert space is

$$HS = \left\{ f : \frac{\partial f(x)}{\partial x} \text{ exists and } f \to 0 \text{ as } x \to \pm\infty \right\}$$

Solution: The objective is to find O^+ such that

$$\langle f | \hat{O} g \rangle = \langle \hat{O}^+ f | g \rangle$$

Start with the quantity on the left

$$\langle f | \hat{O} g \rangle = \int_{-\infty}^{\infty} dx\, f^*(x) \hat{O} g(x) = \int_{-\infty}^{\infty} dx\, f^*(x) \frac{\partial}{\partial x} g(x)$$

The procedure is usually to integrate by parts:

$$\langle f | \hat{O} g \rangle = f^*(x) g(x) \Big]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} dx \frac{\partial f^*(x)}{\partial x} g(x)$$

Note $f^*(\infty) g(\infty) - f^*(-\infty) g(-\infty) = 0$ by assumption which shows the importance of the boundary conditions. Next move the minus sign and partial derivative under the complex conjugate to find

$$\langle f | \hat{O} g \rangle = \int_{-\infty}^{\infty} dx \left[-\frac{\partial f(x)}{\partial x} \right]^* g(x) = \langle \hat{O}^+ f | g \rangle$$

Note everything inside the bra $\langle\ |$ is put under the complex conjugate $(\ )^*$ in the integral.

The operator O^+ is therefore seen to be $-\frac{\partial}{\partial x}$ or $\left(\frac{\partial}{\partial x} \right)^+ = -\frac{\partial}{\partial x}$.

Example 2: The operator

$$\hat{O}^+ = \left(\frac{1}{i} \frac{\partial}{\partial x} \right)^+$$

where $i = \sqrt{-1}$ can be shown to be self-adjoint. It represents the momentum of a particle in quantum theory

$$\hat{p} = \frac{\hbar}{i} \frac{\partial}{\partial x}$$

which is a physical observable.

Topic 2.5.2: Adjoint and Self-Adjoint Matrices

If T is an operator and $|v\rangle$, $|w\rangle$ are two vectors in the Hilbert space, and $\{|a\rangle\}$ is the basis set for the Hilbert space then

$$\langle w|T|v\rangle = \langle w|1T1|v\rangle = \sum_{ab}\langle w|a\rangle\langle a|T|b\rangle\langle b|v\rangle$$

$$= \sum_{ab}\langle a|w\rangle^{+}\langle a|T|b\rangle\langle b|v\rangle$$

where $|a\rangle$ and $|b\rangle$ are two basis vectors. The components of $|v\rangle$ and $|w\rangle$ are the collection of complex numbers $\langle a|w\rangle$ and $\langle b|v\rangle$ which can be arranged as the column vectors

$$\underline{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \end{bmatrix} \quad \text{and} \quad \underline{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \end{bmatrix}$$

The quantity $\langle w|T|v\rangle$ can be written as

$$\langle w|T|v\rangle = \underline{w}^{+}\underline{T}\underline{v}$$

where

$$\underline{T} = \begin{bmatrix} T_{11} & T_{12} & \cdots & T_{1n} \\ T_{21} & T_{22} & \cdots & T_{2n} \\ \vdots & \vdots & & \vdots \\ T_{n1} & T_{n2} & \cdots & T_{nn} \end{bmatrix}$$

The adjoint of the matrix (for example)

$$\underline{T} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad \text{is} \qquad \underline{T}^{+} = \begin{bmatrix} a^* & c^* \\ b^* & d^* \end{bmatrix}$$

which can also be written as $(\underline{T})^{T^*}$ where "T" in the superscript refers to the transpose operation. Also,

$$\underline{w}^{+} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \end{bmatrix}^{+} = \begin{bmatrix} w_1^* & w_2^* & \cdots \end{bmatrix}$$

These results can be demonstrated from previous definitions if desired.

Topic 2.5.3: Eigenvectors and Eigenvalues of Hermitian Operators

There are two important theorems. Assume H is a Hermitian operator defined on the Hilbert space V. Assume that for each eigenvalue E_n there is an eigenfunction $|\phi_n\rangle$. It is customary to label the eigenfunction by the eigenvalue or by the eigenvalue number as

$$|\phi_n\rangle = |E_n\rangle = |n\rangle$$

Usually, the eigenvalues are listed in order of increasing value

$$E_1 < E_2 < \ldots$$

This last point is not essential to the next theorems. The condition of non-degenerate eigenvalues means that for a given eigenvalue, there is only one eigenvector. The eigenvalues are "degenerate" if for a given eigenvalue, there is more than one eigenvector.

non – degenerate	degenerate
$E_1 \leftrightarrow \|E_1\rangle$	$E_1 \leftrightarrow \|E_1\rangle$
$\vdots$	$E_2 \leftrightarrow \|E_2 1\rangle, \|E_2 2\rangle$
$E_n \leftrightarrow \|E_n\rangle$	$E_3 \leftrightarrow \|E_3\rangle$

It turns out that the degenerate eigenvectors (which means both states have the same "energy" E_n) actually span a two-dimensional space (as an example for E_2 above). Mathematically, a vector is chosen in the subspace spanned by $\{|E_2,1\rangle, |E_2,2\rangle\}$ such that it's orthogonal to the others in the set $\{|E_1\rangle, |E_3\rangle \ldots\}$.

The eigenvalues of a Hermitian operator H are *real* and the eigenvectors corresponding to different eigenvalues are orthogonal. It's a good thing that Hermitian operators have real eigenvalues since observables correspond to Hermitian operators and the results of the observation are the eigenvalues. Energy, for example, will never be observed to be imaginary.

Theorem 1: Self-Adjoint Operators H have real eigenvalues

Proof: Assume $\{|n\rangle\}$ are the eigenvectors corresponding to the eigenvalues $\{E_n\}$; i.e., the eigenvector equation is $H|n\rangle = E_n|n\rangle$. Consider

$$\langle n|H|n\rangle = \langle n|E_n|n\rangle = E_n\langle n|n\rangle = E_n$$

which assumes that the eigenvectors are normalized to unity as $\langle n|n\rangle = 1$. So

$$\langle n|H|n\rangle = E_n \tag{2.5.1}$$

take the adjoint of both sides

$$\langle n|H|n\rangle^+ = (E_n)^+$$

Reversing the factors on the left-hand side and changing the "dagger" into a complex conjugate on the right-hand side provides

$$\langle n|H^+|n\rangle = E_n^*$$

The Hermiticity of the operator H , which is $H=H^{+}$, provides

$$\langle n|H|n\rangle = E_n^* \tag{2.5.2}$$

Equations 2.5.1 and 2.5.2 show

$$E_n = E_n^*$$

which means that E_n is real.

Theorem 2: Orthogonal Eigenvectors

If H is Hermitian then the eigenvectors corresponding to different eigenvalues are orthogonal.

Proof: Assume $E_m \neq E_n$ and start with two separate eigenvalue equations

$$\hat{H}|E_m\rangle = E_m|E_m\rangle \qquad \hat{H}|E_n\rangle = E_n|E_n\rangle$$

operate with $\langle E_n|$ operate with $\langle E_m|$

$$\langle E_n|\hat{H}|E_m\rangle = E_m\langle E_n|E_m\rangle \qquad \langle E_m|\hat{H}|E_n\rangle = E_n\langle E_m|E_n\rangle$$

Take adjoint of both sides

$$\langle E_n|\hat{H}|E_m\rangle = E_n\langle E_n|E_m\rangle$$

where the right hand column made use of the Hermiticity of the operator H and the reality of the eigenvalues E_n. Now compare the results of the two columns to find

$$0 = (E_m - E_n)\langle E_n|E_m\rangle$$

The conditions of the theorem assume that $E_m - E_n \neq 0$ and therefore $\langle E_n|E_m\rangle$. The results of the theorems then says that the eigenvectors of a Hermitian form an orthonormal set

$$B = \{|E_n\rangle = |n\rangle\}$$

which is assumed to be complete. Therefore, B is a basis set.

This previous theorem is important because it assures us that Hermitian operators, which correspond to physical observables, have eigenvectors that form a basis for the vector space of all allowed wave functions. Therefore every allowed wave function must be expressible as a linear combination of the eigenvectors/eigenfunctions. The basis set forms the elementary modes for the physical system. When a measurement is made of the physical observable corresponding to the Hermitian operator, the result will always be one of the eigenvalues and the particle will be found in one of the eigenstates. The total wave function collapses to one of the eigenvectors. The modulus-squared of the expansion coefficients for the total wave function are the probabilities of the total wave function collapsing into a particular eigenstate.

Topic 2.5.4: An Algebra of Operators and Commutators

For this topic, consider linear operators that have a domain and range within a single vector space $T: V \to V$ where V has the basis set $B_v = \{|n\rangle: \ n = 1,2,\ldots\}$. There is a matrix for each operator T and a linear isomorphism, which is " 1-1" and "onto", $M: \hat{T} \to \underline{T}$. The linear isomorphism is important because it assures the properties of the operators and the matrices are the same.

Matrices and operators generally do not commute. Linear operators form an algebra which satisfy the properties

(a) A0=0A=0 where 0 is the "zero" operator
(b) AI=IA=A where I is the "unit" operator
(c) A(B+C)=AB+AC where A, B, C are linear operators
(d) A(BC)=(AB)C
(e) aA=Aa where "a" is complex number

Properties a-e require a definition as to when two operators A and B are equal.

Definition Let A,B be to linear operators that map the vector space V into itself. Then A=B if $A|v\rangle = B|v\rangle$ for every vector $|v\rangle$ in the vector space V.

None of the properties a-e refer to the commutation of two (or more) operators. Two operators commute when

AB=BA or AB-BA=0

The quantity AB-BA is represented by the commutator of A and B:

[A,B]=AB-BA

Therefore two operators A and B commute when [A,B]=0.

Example Evaluate $\left[x, \frac{d}{dx}\right] \neq 0$. As an important note, the commutator is itself an *operator* that operates on functions f(x). Therefore

$$\left[x, \frac{d}{dx}\right] f = \left(x\frac{d}{dx} - \frac{d}{dx}x\right) f(x) = x\frac{df}{dx} - \frac{d}{dx}(xf) = x\frac{df}{dx} - \frac{dx}{dx}f - x\frac{df}{dx} = -f \neq 0$$

Notice that the derivative in the commutator operates on everything to the right.

Commutivity and non-commutivity of the operators are of vital importance to our world. It is the basic concept underlying all of quantum mechanics. As will be discussed laterl, if A and B do not commute then A and B have an associated uncertainty relation

$$\sigma_A \sigma_B \geq C > 0$$

where σ is the standard deviation from probability theory.

Some Theorems on Operator Algebra

Let $\hat{A},\hat{B},\hat{C}$ be operators.

0. $\left[\hat{A},\hat{B}\right]=\hat{A}\hat{B}-\hat{B}\hat{A}$
1. $\left[\hat{A},\hat{A}\right]=0$
2. $\left[\hat{A},\hat{B}\right]=-\left[\hat{B},\hat{A}\right]$
3. $\left[\hat{A},\hat{B}+\hat{C}\right]=\left[\hat{A},\hat{B}\right]+\left[\hat{A},\hat{C}\right.$
4. $\left[\hat{A}+\hat{B},\hat{C}\right]=\left[\hat{A},\hat{C}\right]+\left[\hat{B},\hat{C}\right.$
5. $\left[\hat{A},\hat{B}\hat{C}\right]=\left[\hat{A},\hat{B}\right]\hat{C}+\hat{B}\left[\hat{A},\hat{C}\right]$
6. $\left[\hat{A}\hat{B},\hat{C}\right]=\left[\hat{A},\hat{C}\right]\hat{B}+\hat{A}\left[\hat{B},\hat{C}\right.$
7. If $f=f\left(\hat{A}\right)$ then $\left[f\left(\hat{A}\right),\hat{A}\right]=0$

Properties 1 through 6 are easy to prove by expanding the brackets using the definition of commutator. For example, consider property 5

$$\begin{aligned}\left[\hat{A},\hat{B}\right]\hat{C}+\hat{B}\left[\hat{A},\hat{C}\right]&=\left(\hat{A}\hat{B}-\hat{B}\hat{A}\right)\hat{C}+\hat{B}\left(\hat{A}\hat{C}-\hat{C}\hat{A}\right)\\&=\hat{A}\hat{B}\hat{C}-\hat{B}\hat{A}\hat{C}+\hat{B}\hat{A}\hat{C}-\hat{B}\hat{C}\hat{A}\\&=\hat{A}\hat{B}\hat{C}-\hat{B}\hat{C}\hat{A}\\&=\left[\hat{A},\hat{B}\hat{C}\right]\end{aligned}$$

Properties such as Number 7 are usually proved by Taylor expanding the function. The Taylor expansion of a function of an operator has the form

$$f\left(\hat{A}\right)=\sum_n c_n\hat{A}^n$$

where c_n can be a complex number and "n" is a nonnegative integer. Then the commutator can be expanded as

$$\left[f\left(\hat{A}\right),\hat{A}\right]=\left[\sum_n c_n\hat{A}^n,\hat{A}\right]=\sum_n c_n\left[\hat{A}^n,\hat{A}\right]=0$$

Next, examine what happens when two Hermitian operators $\hat{A},\hat{B}$ commute. Each individual Hermitian operator must have a complete set of eigenvectors which means that each Hermitian operator generates a basis set for the vector space. The next theorem shows that if the operators commute $\left[\hat{A},\hat{B}\right]=0$ then the basis set for the vector space can be chosen to be the same as the one for the operator $\hat{A}$ and for the operator $\hat{B}$.

Theorem2.5.4.1: Commuting Hermitian Operators and the Basis Set

If $\hat{A},\hat{B}$ are Hermitian operators that commute $\left[\hat{A},\hat{B}\right]=0$ then there exists, the eigenvectors $|\xi\rangle$ such that $\hat{A}|\xi\rangle=a_\xi|\xi\rangle$ and $\hat{B}|\xi\rangle=b_\xi|\xi\rangle$

Proof: Assume that A has a complete set of eigenvectors. Let $|\xi\rangle$ be the eigenvectors of $\hat{A}$ such that

$$\hat{A}|\xi\rangle=a_\xi|\xi\rangle$$

Further assume that for each a_ξ *there is only one eigenvector* $|\xi\rangle$. Consider

$$\hat{B}\hat{A}|\xi\rangle = \hat{B}a_\xi|\xi\rangle$$

But $BA = AB$ since $[A,B]=0$ and so

$$a_\xi\left(B|\xi\rangle\right) = Ba_\xi|\xi\rangle = BA|\xi\rangle = AB|\xi\rangle = A\left(B|\xi\rangle\right)$$

which means that $B|\xi\rangle$ is an eigenvector of the operator A. Corresponding to the eigenvalue a_ξ. But there is only one eigenvector for each eigenvalue. So

$$|\xi\rangle \sim B|\xi\rangle$$

or, rearranging the previous expression and inserting a constant of proportionality b_ξ, to get

$$B|\xi\rangle = b_\xi|\xi\rangle$$

This is an eigenvector equation for the operator B; the eigenvalue is b_ξ .
-- -- -- QED -- -- --

The eigenvectors can be written as

$$|\xi\rangle = |a_\xi, b_\xi\rangle = |a,b\rangle = |a\rangle|b\rangle$$

so long as we keep track of which eigenvector goes with which operator.

$$\hat{A}|a\rangle|b\rangle = \left(\hat{A}|a\rangle\right)|b\rangle = a|a\rangle|b\rangle$$
$$\hat{B}|a\rangle|b\rangle = |a\rangle\left(\hat{B}|b\rangle\right) = b|a\rangle|b\rangle$$

It should be clear that the set $\{|a\rangle|b\rangle\}$ forms a basis set for the direct product space.

$$\{|a\rangle|b\rangle\} = \{|a\rangle\}\otimes\{|b\rangle\}$$

where the space spanned by the eigenvectors of A and B are $B_A = \{|a\rangle\}$ and $B_B = \{|b\rangle\}$. Notice that the object $|\xi\rangle = |a\rangle|b\rangle$ is considered to be a single object. Commuting operators refer to "different" vector spaces. If an operator has the form O=AB and [A,B]=0 than the matrix of the operator O can be decomposed.

Theorem2.5.4.2: Common eigenvectors correspond to commuting operators

As an inverse to Theorem 2.5.4.1, if the operators $\hat{A},\hat{B}$ have a complete set of eigenvectors in common then [A,B]=0.

Proof: Let $|v\rangle$ be an element of the direct product space of the eigenvectors for the operators $\hat{A},\hat{B}$ so that it can be expanded as

$$|v\rangle = \sum_{ab}\beta_{ab}|a\,b\rangle$$

then

$$AB|v\rangle = \sum_{ab}\beta_{ab}AB|ab\rangle = \sum_{ab}\beta_{ab}Ab|ab\rangle\sum_{ab}\beta_{ab}ba|ab\rangle$$

$$= \sum_{ab}\beta_{ab}aB|ab\rangle = \sum_{ab}\beta_{ab}Ba|ab\rangle\sum_{ab}\beta_{ab}BA|ab\rangle$$

$$= BA|v\rangle$$

This is true for all factors in the vector space and so

$$\hat{A}\hat{B} = \hat{B}\hat{A}$$

-- -- --

Section 2.6: Constraints and Generalized Coordinates

Constraints represent a-priori knowledge of a physical system. They represent a reduction in the total number of degrees of freedom available to the system. For example, Figure 2.6.1 shows rigid rods fixing the distance between a collection of masses so as to reduce the number of degrees of freedom; however, the three masses as a unit can still move. As another example, walls of a container also impose constraints on a system. For quantum theory, constraints are only used to model a system since, in actuality, small particles experience forces and not constraints. For example, an atom is held in place in a lattice by electrostatic forces and not by rigid rods. Evidently, constraints are mostly important for macroscopic classical systems.

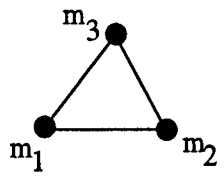


Figure 2.6.1: Three mass connect by rigid rods.

A generalized set of coordinates $B_q=\{q_1,q_2,...,q_k\}$ describe the position of N particles. The q_i are independent of each other without constraints. N particles are normally described by the position vectors

$$\vec{r}_1 = \vec{r}_1(q_1,...,q_k,t)$$
$$\vdots$$
$$\vec{r}_N = \vec{r}_N(q_1,...,q_k,t)$$

For example, the $\{q_i\}$ might be the spherical coordinates. Without constraints, N particles have 3N degrees of freedom so that k=3N. Constraints reduce the degrees of freedom so that k<3N; that is, the constraints eliminate 3N-k degrees of freedom. Configuration space consists of the collection of the "k" generalized coordinates $\{q_1,q_2,...,q_k\}$ where each coordinate can take a range of values. These generalized coordinates are particularly important for the Lagrange formulation of dynamics. The generalized velocities are

$$\left\{\dot{q}_1,\dot{q}_2,...,\dot{q}_k\right\}$$

They are not treated as independent of the generalized coordinates for the Lagrange formulation. That is, the variations δq, $\delta\dot{q}$ are not independent. This book assumes that there aren't any constraints.

A system follows a particular curve in phase space as a function of time. Phase space consists of the generalized coordinates and conjugate momentum (to be defined later)

$$\{q_1,q_2,...,q_k,p_1,p_2,...,p_k\}$$

all of which are assumed to be independent of one another. Assigning particular values to the 2k coordinates in phase space specifies the "state of the system". The phase space coordinates are used primarily with the Hamiltonian of the system.

The generalized coordinates discussed so far consitute a discrete set whereby the coordinates are in one-to-one correspondence with a subset of the integers. It is possible for the set to be infinite. A continuous set of coordinates would have elements in 1-1 correspondence with a dense subset of the *real* numbers. The distinction is important for a number of topics especially field theory.

Of importance is the way the generalized coordinates and velocities can be pictured (especially for a field theory). For a physical medium like a crystal lattice, suppose that there are k atoms in number. The top portion of Figure 2.6.2 shows the atoms at equilibrium. There is one atom for each equilibrium position x_i. The atoms can be labeled be either the respective equilibrium position x_i or by the number "i". The bottom portion of the figure shows the atoms displaced along the vertical direction. In this case, the generalized coordinates label the displacement from equilibrium; for 3-D motion, each atom would have 3 generalized coordinates and 3 generalized velocities. For the 1-D case shown, the generalized coordinates can be written equally well in either of two ways

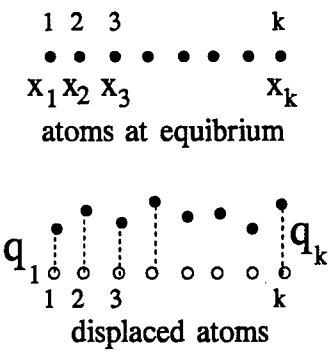


Figure 2.6.2: Example of generalized coordinates for atoms in a lattice.

$$q_i = q(x_i)$$

Mathematically, the displacements just described can be randomly assigned. It's only once the dynamics (Newton's Laws etc) are formulated and applied to the problem that the displacements become correlated. Mathematically, without dynamics (Newton's laws), atom #1 can be moved to position q_1 and atom #2 to position q_2 without there being any reason for choosing those two positions. The notion of independent translations leads to an alternate formulation of Newton's laws.

The Hamilton formulation of dynamics uses phase-space coordinates. These phase-space coordinates are all on the same "footing" meaning that none are more fundamental than the others

$$\{q_1, q_2, \ldots, q_k, p_1, p_2, \ldots, p_k\}$$

For phase space, the variations $\delta q, \delta p$ are considered to be independent of one another. The term *configuration space* is reserved for the "q" coordinates and "phase-space" applies to the full set of 2k coordinates q,p. Essentially, in the absence of dynamics, position and momentum can be arbitrarily assigned to each atom.

Section 2.7: Lagrange's Equation from a Variational Principle

The forces acting on a system of particles control the dynamics of the system. However, forces of constraint are often not known until after the problem is solved. D'Alembert (and Bernoulli) divided the forces into "applied forces" and "constraint forces" and assumed that the virtual work done by forces of constraint is zero (since the forces are assumed to act perpendicular to the direction of motion). The resulting derivation results in the Lagrange formulation of mechanics. Recall that Lagrange's equations provide an alternative formulation of Newton's laws. Maxwell's equations can be most elegantly recast into a Lagrangian. This section reviews the common variational method of obtaining the Lagrangian and Lagrange's equations.

Lagrange's Equation is obtained from Hamilton's principle for conservative systems. The method is particularly easy to generalize for systems consisting of continuous sets of coordinates (i.e., field theory). Of all the possible paths in configuration space that a system could follow between two fixed points $1=(q_1^{(1)},q_2^{(1)},...,q_k^{(1)})$ and $2=(q_1^{(2)},q_2^{(2)},...,q_k^{(2)})$, the path that it actually follows makes the following intergral an extremum (either minimum or maximum).

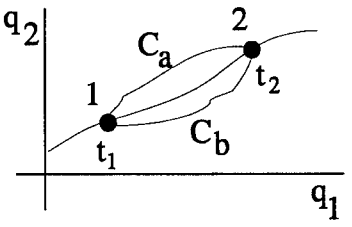


Figure 2.7.1: Three paths connecting fixed end points.

$$I=\int_1^2 dt\, L(q_1,q_2,...,q_k,\dot{q}_1,\dot{q}_2,...,\dot{q}_k,t)$$

"L" is the Lagrangian and is a functional of the kinetic energy "T" and potential energy "V" as in L=T-V. As a note, it is possible to generalize the formulation for variable end-points. To minimize the notation, let $q_i,\dot{q}_i$ represent the entire collection of points in $\{q_1,q_2,...,q_k,\dot{q}_1,\dot{q}_2,...,\dot{q}_k\}$.

To find the extremum of

$$I=\int_1^2 dt\, L(q_i,\dot{q}_i,t)$$

define a new path in configuration space for each generalized coordinate q_i by

$$q_i(t,\alpha_i)=q_i(t,0)+\alpha_i\eta_i(t)=q_i(t,0)+\delta q_i$$

where the time "t" parameterizes the curve in configuration space, α_i are parameters and η_i are the changes in the path. This equation can be differentiated with respect to time to obtain the similar version for $\dot{q}_i$. Assume that $\alpha_i=0$ corresponds to the path $q_i(t,0)$ that extremizes the integral "I". Similarly, $\alpha_i=1$ might correspond to curve C_b in the figure. The exact form of $\eta_i(t)$ is immaterial except that $\eta_i(t_1)=\eta_i(t_2)=0$ for i=1,...,k. The extremum is defined as

$$0=\left(\frac{\partial I}{\partial\alpha_i}\right)_{\alpha_i=0}=\int_1^2 dt\left(\frac{\partial L(q_i,\dot{q}_i,t)}{\partial\alpha_i}\right)_{\alpha_i=0}=\int_1^2 dt\sum_i\left[\frac{\partial L(q_i,\dot{q}_i,t)}{\partial q_i}\left(\frac{\partial q_i}{\partial\alpha_i}\right)_{\alpha_i=0}+\frac{\partial L(q_i,\dot{q}_i,t)}{\partial\dot{q}_i}\left(\frac{\partial\dot{q}_i}{\partial\alpha_i}\right)_{\alpha_i=0}\right]$$

Lagranges Equations obtain by using

$$\eta_i = \left(\frac{\partial q_i}{\partial \alpha_i}\right)_{\alpha_i=0} \qquad \dot{\eta}_i = \left(\frac{\partial \dot{q}_i}{\partial \alpha_i}\right)_{\alpha_i=0}$$

and partial integrating, and assuming all the η_i are independent. However, it is customary to work with the parameters

$$\delta q_i = \alpha_i \eta_i$$

The extremum condition takes the form

$$0 = \delta I = \int_1^2 dt \sum_i \left[\frac{\partial L(q_i, \dot{q}_i, t)}{\partial q_i}\delta q_i + \frac{\partial L(q_i, \dot{q}_i, t)}{\partial \dot{q}_i}\delta \dot{q}_i\right]$$

Partial integrate the second term using the fact that $\delta q_i(t_1) = \delta q_i(t_2) = 0$ to find

$$0 = \delta I = \int_1^2 dt \sum_i \left[\frac{\partial L(q_i, \dot{q}_i, t)}{\partial q_i} + \frac{d}{dt}\frac{\partial L(q_i, \dot{q}_i, t)}{\partial \dot{q}_i}\right]\delta q_i$$

The small variations δq_i are assumed to be independent so that

$$\frac{\partial L}{\partial q_i} + \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = 0 \qquad i = 1,2,...$$

where L=T-V.

The canonical momentum is defined to be

$$p_i = \frac{\partial L}{\partial \dot{q}_i}$$

p_i is the momentum conjugate to the coordinate q_i. The canonical momentum is not always the same as the usual momentum "mv" for a particle. The canonical momentum for an EM field interacting with a particle consists of the particle and field momentum.

Example: Single particle of mass "m" constrained to move vertically along the "y" direction and acted upon by the graviational force $F = -mg$

$$T = \frac{1}{2}m(\dot{y})^2 \qquad V = mgy \qquad L = T - V = \frac{1}{2}m(\dot{y})^2 - mgy$$

Lagrange's equation

$$\frac{\partial L}{\partial y} - \frac{d}{dt}\frac{\partial L}{\partial \dot{y}} = 0$$

gives Newton's second law for a gravitational force

$$-mg - m\dot{y} = 0$$

where the derivatives

$$\frac{\partial \dot{y}}{\partial y} = 0 = \frac{\partial y}{\partial \dot{y}}$$

since "y" and "$\dot{y}$" are taken to be independent.

Section 2.8: The Hamiltonian

The Hamiltonian is the total energy of a system (under certain conditions). The Hamiltonian and Lagrangian formalism provide the description and dynamics quantum theory. The Hamiltonian provides deep insight into the structure of mechanics.

Consider a closed, conservative system. The Lagrangian L is not an explicit function of time. A closed system is one for which the total energy and the total number of particles are constant. A conservative system is defined to have all of the forces derivable from a potential function. There aren't any equations of constraint for quantum mechanics and field theory. Differentiating the Lagrangian provides

$$\frac{dL}{dt} = \sum_i \left(\frac{\partial L}{\partial q_i} \frac{dq_i}{dt} + \frac{\partial L}{\partial \dot{q}_i} \frac{d\dot{q}_i}{dt} \right) + \frac{\partial L}{\partial t}$$

The last term is zero by assumption

$$\frac{\partial L}{\partial t} = 0$$

Substitute Lagrange's equation

$$\frac{\partial L}{\partial q_i} = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i}$$

to find

$$\frac{dL}{dt} = \sum_i \left[\left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) \dot{q}_i + \frac{\partial L}{\partial \dot{q}_i} \frac{d\dot{q}_i}{dt} \right] = \sum_i \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) \qquad (2.8.1)$$

Recall the definition for the generalized coordinates

$$p_i = \frac{\partial L}{\partial \dot{q}_i}$$

Equation 2.8.1 becomes

$$\frac{d}{dt} \left[\sum_i \dot{q}_i p_i - L \right] = 0$$

The Hamiltonian "H" is defined to be

$$H = \sum_i \dot{q}_i p_i - L$$

which is the total energy of the system in this case. *Important point:* H is considered to be a function of q_i, p_i whereas L is a function of $q_i, \dot{q}_i$.

Hamilton's canonical equations can now be found. Using the fact that H is a function of q_i, p_i (whereas L is a function of $q_i, \dot{q}_i$) and using the definition of the conjugate momentum, the following partial derivatives of H hold

$$\frac{\partial H}{\partial p_j} = \frac{\partial}{\partial p_j} \left[\sum_i \dot{q}_i p_i - L \right] = \dot{q}_j - \frac{\partial L}{\partial p_j} = \dot{q}_j$$

and

$$\frac{\partial H}{\partial q_j} = \frac{\partial}{\partial q_j} \left[\sum_i \dot{q}_i p_i - L \right] = 0 - \frac{\partial L}{\partial q_j} = -\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} = -\frac{d}{dt} p_j = -\dot{p}_j$$

Example: For the example in Section 2.7, find H and $\dot{q}_i, \dot{p}_i$.
Solution: We had

$$L = T - V = \frac{1}{2}m(\dot{y})^2 - mgy$$

H must be written as a function of the coordinate and its conjugate momentum. The relation for the canonical momentum for the Lagrangian

$$p = \frac{\partial L}{\partial \dot{y}} = m\dot{y}$$

allows "H" to be written as

$$H = \dot{y}p - L = \frac{p}{m}p - \left[\frac{1}{2}m\left(\frac{p}{m}\right)^2 - mgy\right] = \frac{p^2}{2m} + mgy$$

and then

$$\dot{y} = \frac{\partial H}{\partial p} = \frac{p}{m} \qquad \dot{p} = -\frac{\partial H}{\partial y} = mg$$

The Hamiltonian H can be seen to be the sum of the kinetic and potential energy T+V by calculating

$$H = \sum_i \dot{q}_i p_i - L$$

with L=T-V and using a general quadratic form for the kinetic energy

$$T = \sum_{i,j} a_{ij}\dot{q}_i\dot{q}_j \qquad \text{where } a_{ij}=a_{ji}$$

The canonical momentum is

$$p_m = \frac{\partial L}{\partial \dot{q}_m} = 2\sum_i a_{im}\dot{q}_i$$

Therefore,

$$H = \sum_m \dot{q}_m p_m - L = \sum_m \dot{q}_m 2\sum_i a_{im}\dot{q}_i - (T-V) = 2\sum_{mi} a_{im}\dot{q}_i\dot{q}_m - T + V$$
$$= 2T - T + V$$
$$= T + V$$

Section 2.9: Poisson Brackets

The Hamiltonian is the primary quantity of interest for quantum theory. The specification of a quantum mechanical Hamiltonian follows several steps

(1) Determine the classical Hamiltonian
(2) Substitute operators for the classical dynamical variables (e.g., p's and q's)
(3) Specify the commutation relations between those dynamical variables

The commutation relations in quantum mechanics are similar in function to the Poisson brackets in classical mechanics. The commuation relations and Poisson brackets determine the evolution of the dynamical variables. In the quantum theory, operators replace the classical dynamical variables. In fact, the Heisenberg quantum picture is the closest analogue to classical mechanics because the operators carry the system dynamics. In quantum theory, the commutation relations give time derivatives of operators. A commutator is defined by $\left[\hat{A},\hat{B}\right]=\hat{A}\hat{B}-\hat{B}\hat{A}$ where $\hat{A},\hat{B}$ are operators. The classical theory is given below but notice that the Poisson bracket involves partial derivatives whereas the quantum mechanical commutator does not.

Definition: Let $A=A(q_i,p_i)$ $B=B(q_i,p_i)$ be two differentiable functions of the generalized coordinates and momentum. The Poisson brackets are defined by

$$[A,B]=\sum_i\left[\frac{\partial A}{\partial q_i}\frac{\partial B}{\partial p_i}-\frac{\partial B}{\partial q_i}\frac{\partial A}{\partial p_i}\right]$$

Sometimes the brackets are subscripted with p,q to distinguish which phase space variables are used

$$[A,B]=[A,B]_{q,p}$$

Using the basic definition of Poisson brackets, some basic properties can be proved.

(1) Let A, B be functions of the phase space coordinates q,p and let c be a number then

$$[A,A]=0 \qquad [A,B]=-[B,A] \qquad [A,c]=0$$

(2) Let A, B, C be differentiable functions of the phase space coordinates q,p then

$$[A+B,C]=[A,C]+[B,C] \qquad [A,BC]=[A,B]C+B[A,C]$$

(3) The time evolution of the dynamical variable A (for example) can be calculated by

$$\frac{dA}{dt}=[A,H]+\frac{\partial A}{\partial t}$$

Proof:

$$\frac{dA}{dt}=\sum_i\left[\frac{\partial A}{\partial q_i}\frac{dq_i}{dt}+\frac{\partial A}{\partial p_i}\frac{dp_i}{dt}\right]+\frac{\partial A}{\partial t}$$

Where the partial with respect to time handles the case that A is an explicit function of time. Substitute the two relations for the rate of change of position and momentum

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i} \qquad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}$$

so that the Poisson brackets become

$$\frac{dA}{dt} = \sum_i \left[\frac{\partial A}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial A}{\partial p_i}\frac{\partial H}{\partial q_i} \right] + \frac{\partial A}{\partial t} = [A,H] + \frac{\partial A}{\partial t}$$

Although the order of the quantities A,H does not matter in classical theory, the order must be maintained in quantum theory. In quantum theory, the order of two operators can only be switched by using the commutation relations.

(4) $\dot{q}_m = [q_m, H] \qquad \dot{p}_m = [p_m, H]$

Proof: Consider the first one for example

$$[q_m, H] = \sum_i \left[\frac{\partial q_m}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial q_m}{\partial p_i}\frac{\partial H}{\partial q_i} \right] = \sum_i \left[\delta_{im}\frac{\partial H}{\partial p_i} - 0\frac{\partial H}{\partial q_i} \right] = \frac{\partial H}{\partial p_m} = \dot{q}_m$$

(5) $[q_i, q_j] = 0 \qquad [p_i, p_j] = 0 \qquad [q_i, p_j] = \delta_{ij}$

These properties are all very similar to those that arise in the quantum theory.

Section 2.10 Introductory Quantum Mechanics Review

Quantum mechanics forms a cornerstone for physics, engineering and chemistry. Quantum Theory has been developing since the turn of the century. It has found significant applications in engineering due to the development of semiconductor devices, the laser and nanometer scale devices. During the 1980s, the fabrication and materials growth technology developed far enough to provide the ability to (1) produce quantum-well devices (such as quantum well lasers) and (2) engineer the optical and electrical properties of materials (band-gap engineering). The following few sections review a few of the basic concepts in quantum mechanics.

Section 2.11 discusses the connection between quantum mechanics and the linear algebra. It also discusses the origin of the Heisenberg uncertainty relation. Section 2.11 reiterates some basic theory and applies it to some basic examples including a particle in a well and quantum tunneling. Section 2.13 continues with an important example for the harmonic oscillator and discusses raising and lowering operators. Section 2.12 discusses the Schrodinger, Heisenberg, and interaction representations. Section 2.14 presents the time independent perturbation theory.

Topic 2.10.1: The Relation Between Quantum Mechanics and Linear Algebra

Hermitian operators represent physically observable quantities such as energy, momentum, and electric field. The act of observing these quantities is equivalent to applying the operators to vectors in Hilbert space. There is however, one proviso concerning the collapse of the wavefunction as discussed in topic 2.10.5. The vectors represent the state of the particle (or particles). Every physically possible state of the system must be represented by one of the vectors in the Hilbert space. Function space contains wavefunctions, which describe the particles, as the resident vectors. The physical meaning of the word "state" needs to be defined.

Another important issue concerns the origin of particle dynamics. In the classical sciences and engineering, variables such as energy and momentum depend on time. The most natural extension to quantum theory is to have the corresponding Hermitian operators depend on time, which is called the Heisenberg representation. In the Heisenberg representation, the wavefunctions are independent of time and serve as a type of stage for observation. The result of an observation depends on the time that the observation is made. In the Schrodinger representation, the wavefunctions depend on time very similar to the wavefunctions in classical optics; however, the operators are time independent. There is also an intermediate case where the operators carry a trivial time-dependence and the wavefunctions contain the time response to a "forcing function"; this is the so-called interaction representation.

The discussion addresses issues on (1) how basis vectors differ from other vectors, (2) the meaning of superposition, (3) the physical meaning of the expansion coefficients of a general vector in a Hilbert space, (4) a picture of the time dependent wavefunction (5) observables that cannot be simultaneously observed, and (6) the collapse of the wavefunction.

Topic 2.10.2: The Eigenstates

The eigenstates of a Hermitian operator, which correspond to an observable, are the most fundamental states for the particle or system. Every possible fundamental motion of a particle must be observable. This requires that each fundamental mode must be represented as a basis vector. The basis set must be complete to allow detection of the fundamental physical modes.

$$\hat{H}|\phi_n\rangle = E_n|\phi_n\rangle$$

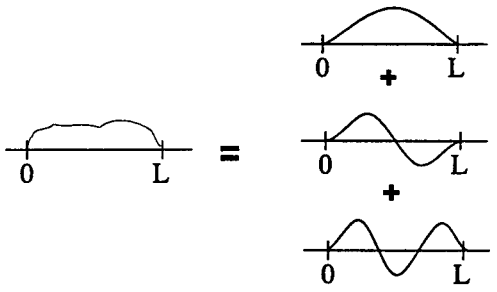


Figure 2.10.1: A classical wave on a string is decomposed into the basic modes (i.e., the basis vectors).

For a classical system, the fundamental modes might be the sine and cosines for a wave on a string; the sines and cosines correspond to the basis vectors. Fourier analysis then provides that a general classical wave on the string is a sum over all the fundamental modes.

The general wavefunction is a linear combination of the basis set according to the method of Fourier series.

$$|\psi(t)\rangle = \sum_n \beta_n(t)|\phi_n\rangle$$

or, equivalently

$$\psi(x,t) = \sum_n \beta_n(t)\langle x|\phi_n\rangle = \sum_n \beta_n(t)\phi_n(x)$$

The basic modes are independent of time. The time-dependence of the vibrational motion is contained in the expansion coefficients $\beta_n(t)$. The basis set originates as the set of eigenvectors for the time-independent wave equation using separation of variables. The total wave function is made up of the basis set.

A difference between the Fourier analysis for the motion of the classical string and that for quantum mechanics is the idea of the collapse of the wave function. Using a string analogy, the collapse of the wave function $|\psi(t)\rangle$ means that a measurement of some observable such as total energy or speed causes the wave function $|\psi(t)\rangle$ to suddenly collapse to one of the basis vectors $|\phi_n\rangle$. In other words, a measurement of the wave on the string (in the quantum sense) would cause the wave to suddenly become a perfectly defined sine wave!

In the classical mechanics of point particles, the state of the particle at a particular time is specified by its position and speed (i.e., technically, the position and momentum). For an extended body (not point particles), the state might be specified by position and momentum of the center of mass and the angular momentum. In optics, the basic states (i.e., modes) are specified by the polarization, wavelength and propagation vector. Notice that the amplitude was not included in the list even though it needs to be specified. The list contains the parameters that describe the basic opitcal mode. The number of quanta occupying the mode gives information on the amplitude. The optical modes are the primary quantity of interest; they are eigenvectors for the time-independent wave

equation. These basic modes are usually sines and cosines for a cavity or they might be travelling waves (without the $e^{i\omega t}$ factor) for free space.

In quantum theory, the basic state (i.e., basis state) of a particle or system is specified by listing the observable properties. The particle might have a certain energy, momentum, angular momentum, polarization etc. Knowing the value of all observable properties is equivalent to knowing the basis state of the particle or system. Each physical "observable" corresponds to a Hermitian operator $\hat{O}_i$ which induces a preferred basis set for the respective Hilbert space V_i (i.e., the set of eigenvectors is the "preferred" basis set). The multiplicity of possible observables means that a single particle can "reside" in many Hilbert spaces at the same time since there can be a Hilbert space V_i for each operator $\hat{O}_i$ (some operators might operate within the same space). The particle can therefore resides in the direct product space given by

$$V = V_1 \otimes V_2 \otimes ...$$

where V_1 might describe the energy, V_2 might describe the momentum and so on. The basis set for the direct product space consists of the combination of the basis vectors for the individual spaces such as

$$|\Psi\rangle = |\phi, \eta, ...\rangle = |\phi\rangle|\eta\rangle...$$

where assume, for example, that the space spanned by $\{|\phi\rangle\}$ refers to the energy content and $\{|\eta\rangle\}$ refers to polarization. Then the polarization associated with a particle might be found as

$$\hat{P}|\Psi\rangle = |\phi\rangle\hat{P}|\eta\rangle = |\phi\rangle p|\eta\rangle = p|\Psi\rangle$$

where "p" is the result of the measurment.

The basis states are labeled by the eigenvalues of the commuting Hermitian operators. For example, $|E_i, p_j\rangle$ is the state of the particle with energy E_i and momentum p_j assuming of course that the Hamiltonian and momentum commute.

Topic 2.10.3: Superposition of Basis States and Probability

A particle can "occupy" a state

$$|v\rangle = \sum_n \beta_n(t)|\phi_n\rangle \qquad (2.10.1)$$

where $\{|\phi_n\rangle\}$ are the basis states for the Hilbert space. Often, the basis set consists of the energy eigenvectors. Recall that the basis vectors are normalized to one and that

$$\langle\phi_m|\phi_n\rangle = \delta_{mn}$$

Wavefunctions are always normalized including the superposed wavefunction in Equation 2.10.1

$$\langle v|v\rangle = 1$$

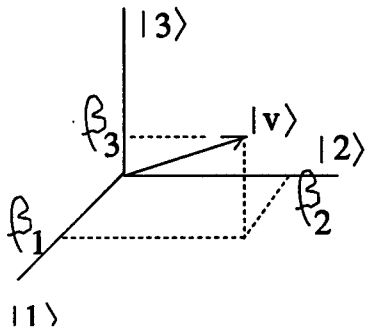


Figure 2.10.2: The decomposition of the function $|v\rangle$ into basis functions.

The normalization of the total wavefunction allows for the interpretation of the expansion coefficients $\beta_n = \langle n|v\rangle$ probability amplitudes. The probability of finding the system in state "n" (i.e., the particle in state $|n\rangle = |\phi_n\rangle$) is given by

$$P(n) = |\beta_n|^2 = |\langle n|v\rangle|^2$$

A probability function must satisfy certain conditions, two of which are that

$$P(n) \geq 0 \text{ and } \sum_n P(n) = 1$$

The modulus of a complex number is always greater than or equal to zero. It's useful to heck the summation property.

$$1 = \langle v|v\rangle = \sum_m \sum_n \beta_m^* \beta_n \langle \phi_m | \phi_n \rangle = \sum_n |\beta_n|^2 = \sum_n P(n)$$

The expansion condition for all wavefunctions provides that the summation of all probabilities equals one. Even though the coefficients β_n depend on time, their sum still gives the value of one.

Topic 2.10.4: Motion of the Wavefunction

The Schrodinger Equation provides the dynamics of the wavefunction. Consider the time dependence of the wavefunctions. The wavefunctions move in Hilbert space since the coefficients β_n depend on time. Notice that the wavefunction does not move out of the given Hilbert space, which is a result of the fact that the eigenvectors form a complete set. Also, because all wavefunctions are normalized to one, the operator "$\hat{u}$" that moves the wavefunction $|\psi\rangle = |\psi(t)\rangle$ in time according to

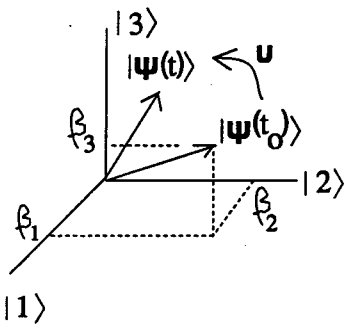


Figure 2.10.3: Motion of the wavefunction under the action of the evolution operator.

$$|\psi(t)\rangle = \hat{u}(t - t_o)|\psi(t_o)\rangle$$

must be a unitary operator. The coefficients depend on time and so do the probabilities $P(n) = |\langle n|v(t)\rangle|^2 = |\beta_n(t)|^2$

Rather than adding probabilities, it is quantum mechanically correct to add amplitudes (i.e., wavefunctions). The procedure is very similar to adding electric fields rather than power (i.e., the square of the electric fields).

Topic 2.10.5: Observables that Cannot be Simultaneously Observed.

This section stipulates that a state of a particle or system is specified by the values of possible observations. Position and momentum are both observables but they cannot be observed at the same time with infinite precision. These observables correspond to Hermitian operators that do not commute. As discussed in the sections on linear algebra, there is no common basis set for the two operators.

Topic 2.10.6: Collapse of the Wavefunction

When an observation of a particular property is made, the particle wavefunction collapses from a sum over all basis vectors to just one of the basis vectors. For example, consider the energy operator and suppose that the system consists of an electron in an infinitely deep well. The energy state of the system must be a vector in the Hilbert space spanned by the states.

$$\left\{ \langle x|\phi_n\rangle = \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi x}{L}\right) \quad n = 1,2,... \right\}$$

so that

$$|v\rangle = \sum_n \beta_n(t)|\phi_n\rangle$$

might represent the wavefunction of the electron. The coefficients are time dependent in general, but this discussion considers t=0 and set $\beta_n(t)=\beta_n$. Previous sections have said that the act of observing the energy is equivalent to applying the Hamiltonian operator to *certain* states. Applying the Hamiltonian $\hat{H}$ to the vector $|v\rangle$ gives

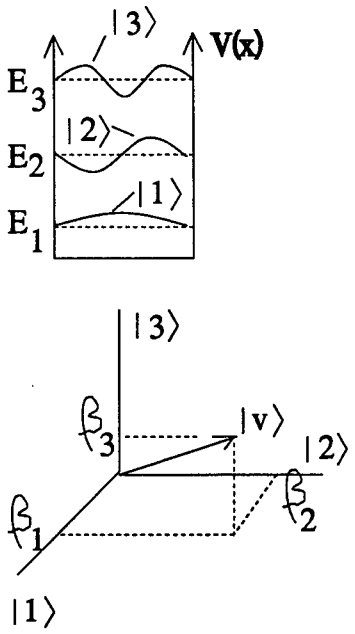


Figure 2.10.4: Wavefunctions for the infinitely deep well.

$$\hat{H}|v\rangle = \sum_n \beta_n(t)\hat{H}|\phi_n\rangle = \sum_n \beta_n(t)E_n|\phi_n\rangle \qquad (2.10.2)$$

So what is the result of the observation? While the last equation is mathematically correct, it does not accurately model the "act of observing!" The observation of the superposition must disturb the total wavefunction and cause a collapse to one of the eigenstates of the Hamiltonian! The process of observing a particle must therefore involve a projection operator. Suppose that the observation causes the wave function to collapse to state 2 (of course it could also collapse to states 1 or 3 with non-zero probably). The act of observing the energy state should include a projection operator $\hat{P}_2 = \frac{1}{\beta_2}\langle\phi_2|$ (not to be confused with the momentum operator). The act of observing the energy can be represented by $\hat{P}_2\hat{H}$ where $\hat{P}_2$ includes a normalization constant of $1/\beta_2$. The results of the observation is then

$$\hat{P}_2\hat{H}|v\rangle = \left\{\frac{1}{\beta_2}\langle\phi_2|\right\}\hat{H}|v\rangle = \sum_n \beta_n(t)\frac{1}{\beta_2}\langle\phi_2|\hat{H}|\phi_n\rangle = E_2$$

It is not possible to know a-priori into which state the wavefunction $|v\rangle$ will collapse. Only the probability of collapsing into a particular state can be given. The probability of it collapsing into state $|n\rangle$ is $|\beta_n|^2 = \beta_n^*\beta_n = |\langle\phi_n|v\rangle|^2$, which is obviously related to the expansion coefficients $\beta_n(t)$. People generally make a distinction between an undisturbed and a disturbed wavefunction. The undisturbed wavefunction can be used as in Equation 2.10.2 and follows the dynamics embedded in Schrodinger's equation. The disturbed wavefunction has collapsed (or made a transistion) to one of the basis states at some point

in time. Without a randomizing process, the wavefunction stays in the eigenstate once it has collapsed to that state. For the eigenstates, the process of observation gives

$$\hat{H}|\phi_n\rangle = E_n|\phi_n\rangle$$

The principal of quantum mechanics, as distinct from the mathematics, is that the act of observing a particular observable causes the wave function to collapse to one of the operators eigenstates. The result of the observation is the eigenvalue E_n. The expansion coefficients for a general state give the probability of observing the corresponding energies (i.e., the probability that the state vector will collapse to the corresponding eigenstate).

Section 2.11: Some Basic Elements of Quantum Mechanics

Recall that elementary particles such as electrons, holes, photons, or phonons can be pictured as waves but retain particle characteristics through momentum and energy relations. This section sometimes refers to massive particles as electrons even though, obviously, not all massive particles are electrons (this is just a bad habit). And equally as bad, it sometimes refer to massless particles as photons. The momentum and energy relations are given by

$$p = \hbar k \qquad E = \hbar\omega$$

where

$$\hbar = \frac{h}{2\pi}$$

and "h" is Planck's constant. For both massive and massless particles, the wave vector and angular frequency are

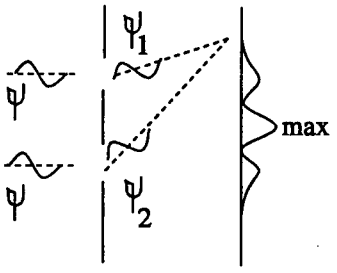
$$k = \frac{2\pi}{\lambda} \qquad \omega = 2\pi\nu$$

where λ and ν are the wavelength and frequency. For massive particles, the momentum is related to the mass "m" and speed "v" by p=mv; consequently, the wavelength of the massive particle is related to the speed and mass by

$$\lambda_o = \frac{h}{mv}$$

More generally however plots of energy E vs. wave vector k give the dispersion curves for the particles; as a result, it is necessary know $\omega(k)$ to find the speed of the wave packet.

The electron wave function is similar to that for the photon and is usually denoted by $\Psi(\vec{r},t)$ where $\vec{r}$ is the position vector and "t" is the time; notice that $\vec{r}$ is independent of time. Light, for example, can be represented by the wave function usually called the electric field (or, more generally, the vector potential). We also note that the square of the electric field is related to the power carried by the light beam; power is really a measure of the number of photons in the beam as we discussed in connection with the photon density. Similarly, the square of the massive particle wave function $\Psi^*(\vec{r},t)\Psi(\vec{r},t)$, is related to the number of particles in a given region of space. In fact, this amplitude is the probability density (probability per volume) for finding a particle in the volume.



As an illustration for picturing an electron (note: a single electron), consider an electron interference experiment. Assume that a *single* electron is incident from the left on to closely spaced slits. The wave function for the electron divides into two coherent pieces denoted by ψ_1 and ψ_2 which traveled toward the screen. The figure shows (at the screen) of a square of the superposition wave function $|\Psi_1 + \Psi_2|^2$ which is the probability of finding an electron at a particular point. The superposition is equivalent to adding electric fields in the classical to slit experiment for light. The interpretation is similar for both the massive and the zero-mass cases.

Previous sections discuss these wave functions as part of a linear vector space. A given wave function can be decomposed into a set of basis functions. Similar to light modes in the cavity, the basis functions are the simplest solutions for a given problem and, upon which, more complicated functions can be constructed. The basis functions are viewed as the time independent solutions to a partial differential equation. Any linear combination of either wave functions or basis functions produces another wave function; however, only the "direction" of that wave function has physical significance. The length of the wave function representing the particle is usually normalized to one. The normalization places restrictions on the expansion coefficients so that, strictly speaking, the collection of wave functions form a manifold in the vector space (and not actually the full vector space). For light, the partial differential equation is the wave equation obtained from Maxwells equations and, for massive particles, the partial differential equation is Schrodinger's equation. All of the basic properties of massive and zero-mass particles are assumed to be contained in the wave functions. The act of observing certain properties of a particle is equivalent to operating on the wave function with the corresponding operator.

All observables are represented by "Hermitian" operators. Any physically measurable quantity is termed an observable. For example, the momentum of a particle is an observable just as it is temperature, electric field, position and everything else that we can measure in a laboratory. As has been discussed, quantum mechanics does not allow the simultaneous observation of all quantities that can be separately observed. For example, position and momentum of a particle cannot *simultaneously* be measured with infinite accuracy for both quantities. This is related to the Heisenberg uncertainty relations.

Hermitian operators $\hat{O}$ correspond to an observable and if "ϕ" is the wave function, then the result of the observation is given by the eigenvector equation

$$\hat{O}\phi = o\phi$$

where "o" is a *real* constant (it is not an operator) and *"o" is the result of the observation.* If for example, "O" is the momentum operator, then "o" is the momentum of the particle when the particle is in state "ϕ". The function "ϕ" is one of the basis vectors which are eigenfunctions of the operator "O". The value "o" is the eigenvalue. For every observable, it is always possible to write an eigenfunction equation. The result of every physical observation must always be an eigenvalue.

Topic 2.11.1: The Schrodinger Wave Equation

It is possible to observe the total energy of a particle or a system of particles. The Hamiltonian $\hat{H}$ is the Hermitian operator that represents the total energy. The Schrodinger wave equation is

$$\hat{H}\Psi = i\hbar\frac{\partial\Psi}{\partial t}$$

The wave function may be an arbitrary function which can be Fourier expanded in terms of the basis set. The energy operator must be specified in terms of other quantities related to the energy of the system. For a single particle, the total energy can be related

to the momentum and potential. The Hamiltonian $\hat{H}$ is an operator and must therefore contain operators.

The usual procedure for finding the quantum mechanical Hamiltonian is to write the classical Hamiltonian (i.e., energy) and then substitute operators for the dynamical variables (i.e., observables). The operators are then required to satisfy commutation relations which accounts for the fact that the corresponding observables might or might not be simultaneously observable (i.e., the Heisenberg uncertainty relations must be satisfied).

The mathematical theory of quantum mechanics admits many different forms for the operators. In the spatial-coordinate representation, the momentum is related to the spatial gradient. In the momentum representation (i.e., Fourier transform space), the momentum operator is written as $\hat{p} = \hbar k$. A single particle has a classical Hamiltonian of the form

$$H = \frac{p^2}{2m} + V(\vec{r})$$

where $p^2/2m$ is the kinetic energy of the particle, V is its potential energy, and the momentum squared is $\vec{p} \cdot \vec{p}$. The quantum mechanical Hamiltonian for a problem with only one degree of freedom (for example, the $\hat{x}$-direction) is found by substituting

$$\hat{P}_x = \frac{\hbar}{i}\frac{\partial}{\partial x}$$

for the classical "p". For three dimensions, the momentum operator is related to the gradient

$$\hat{P} = \frac{\hbar}{i}\nabla$$

The Hamiltonian can be written as

$$\hat{H} = \frac{\hat{P}^2}{2m} + V(\vec{r}) = \frac{1}{2m}\left(\frac{\hbar}{i}\nabla\right)\bullet\left(\frac{\hbar}{i}\nabla\right) + V(\vec{r}) = -\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r})$$

For a single degree of freedom, the commutation relations is

$$\left[x, \hat{P}_x\right] = i\hbar$$

For three degrees of freedom, define the notation

$$x_m = \begin{cases} x & m = 1 \\ y & m = 2 \\ z & m = 3 \end{cases} \qquad \hat{P}_m = \begin{cases} \hat{P}_x & m = 1 \\ \hat{P}_y & m = 2 \\ \hat{P}_z & m = 3 \end{cases}$$

so that the commutation relations can be written as

$$\left[x_m, \hat{P}_n\right] = i\hbar\delta_{mn}$$

Only conjugate variables (i.e., m=n) do not commute. Conjugate variables refer to the same degree of freedom.

Topic 2.11.2: Quick Review of Probability Density and Operator Moments

The probability density function ρ measures the probability per unit length or per unit volume. If ρ(x) dx is the probability of finding a particle in the infinitesimal interval dx centered at the position x, then the probability of finding the particle in the interval [a,b] is given by

$$P(a \le x \le b) = \int_a^b dx\, \rho(x)$$

and averages are calculated as usual

$$\langle f(x) \rangle = \int_{-\infty}^{\infty} dx\, f(x) \rho(x)$$

As is typical for classical probability theory, the density function ρ must be in [0,1] and

$$\int_{-\infty}^{\infty} dx\, \rho(x) = 1$$

The fact that the integral over all space equals unity is a reflection of the fact that the particle must be found somewhere in space (i.e., the total probability is one for finding the particle somewhere). For volume rather than length, the probability density is ρ(x,y,z)

$$P(a \le x \le b, c \le y \le d, e \le z \le f) = \int_a^b \int_c^d \int_e^f dx\, dy\, dz\, \rho(x,y,z) = \int_V dV\, \rho$$

with similar formulas for averages etc.

In quantum theory, the probability density is given by

$$\rho = \Psi^* \Psi = |\Psi|^2$$

and the wave functions are normalized such that

$$\int_V dV |\Psi|^2 = 1$$

where the volume can neither be finite or infinite depending on the physical situation. This chapter will show that integrals similar to the last one are related to inner products and the "length" of a function when viewed as a vector.

Average of an Operator

The correct form of a quantum mechanical average of an observable **O** is

$$\langle \hat{O} \rangle \equiv \langle \Psi | \hat{O} | \Psi \rangle \equiv \int_{-\infty}^{\infty} dx\, \Psi^* \hat{O} \Psi$$

Sometimes the expectation value is written as

$$E(\hat{O}) = \langle \hat{O} \rangle$$

The wave function must be known to calculate the expectation values since the notation does not necessarily make it clear.

The Variance and Standard Deviation

The variance of an operator is defined by

$$\sigma^2 = E\left(\hat{O} - \langle\hat{O}\rangle\right)^2 = E\left(\hat{O}^2 - 2\hat{O}\langle\hat{O}\rangle + \langle\hat{O}\rangle^2\right) = E\left(\hat{O}^2\right) - \langle\hat{O}\rangle^2 = \langle\hat{O}^2\rangle - \langle\hat{O}\rangle^2$$

The standard deviation is

$$\sigma = \sqrt{\langle\hat{O}^2\rangle - \langle\hat{O}\rangle^2}$$

Once again notice that the wavefunction must be known to make the calculations.

Topic 2.11.3: Commutators and the Heisenberg Uncertainty Relations

Elementary theory says that certain observables cannot be measured simultaneously with infinite precision. For example, position and momentum as conjugate variables must satisfy the Heisenberg uncertainty relations

$$\sigma_x \sigma_p \geq \hbar/2$$

or, in other notation, $\Delta x\, \Delta p \geq \hbar/2$. The symbols σ_x, σ_p represent the standard deviation in the position and momentum (corresponding to the x-direction). The standard deviations σ_x, σ_p are not operators since the expectation values have been calculated. The Heisenberg uncertainty relation says that repeated measurements of position and momentum yields a range of values x and p. Using linear algebra, it is possible to show that the Heisenberg uncertainty relation obtains from properties of the operators.

To say that momentum and position cannot be measured precisely at the same time is equivalent to writing

$$\hat{P}_x \hat{x}\Psi \neq \hat{x}\hat{P}_x\Psi$$

or introducing the commutator notation

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$$

we can write

$$[\hat{x}, \hat{P}_x] \neq 0$$

Notice that the commutator is itself an operator and must always operate on a function.

Example: The commutation relation for position and momentum can be evaluated using spatial coordinates.

$$\hat{x} = x \qquad \hat{P}_x = \frac{\hbar}{i}\frac{\partial}{\partial x}$$

and the commutator is evaluated as

$$[x, \hat{P}_x]\Psi = x\frac{\hbar}{i}\frac{\partial}{\partial x}\Psi - \frac{\hbar}{i}\frac{\partial}{\partial x}(x\Psi) = i\hbar\Psi$$

The previous equation is required to hold for every function in the vector space and the commutator must be

$$[\hat{x}, \hat{P}_x] = i\hbar$$

Also, $\qquad [\hat{x}, \hat{x}] = 0 \qquad [\hat{y}, \hat{P}_x] = 0$

-- -- --

The previous example shows that the y-position coordinate commutes with the momentum for the x-coordinate. Commuting operators corresponding to dynamical variables that can be simultaneously and precisely measured. As previously discussed, if two operators $\hat{A},\hat{B}$ commute then there exists a simultaneous set of basis functions $|a,b\rangle = |a\rangle|b\rangle$ such that

$$\hat{A}|a,b\rangle = a|a,b\rangle \quad \text{and} \quad \hat{B}|a,b\rangle = b|a,b\rangle$$

and vice versa. This topic shows that if two operators do not commute then there exists a Heisenberg's uncertainty relation connecting them. As a note, the quantum mechanical Hamiltonian obtains from the classical Hamiltonian by substituting operators for dynamical variables. If we cannot simultaneously and precisely measure both momentum and position, it would seem that the energy of the particle is not well defined!?! The resolution to this apparent contradiction is that the Hamiltonian is well defined for an energy eigenfunctions basis set: the Hamiltonian will be well defined even though momentum and position cannot be simultaneously measured exactly.

The discussion now demonstrates that two noncommuting Hermitian operators must always produce an uncertainty relation.

Theorem: If two operators $\hat{A},\hat{B}$ are Hermitian and satisfy the commutation relation $\left[\hat{A},\hat{B}\right] = i\hat{C}$ then the observed values "a,b" of the operators must satisfy a Heisenberg uncertainty relation of the form $\sigma_a\sigma_b \geq \frac{1}{2}\left|\langle\hat{C}\rangle\right|$.

Proof: Consider the *real, positive number* defined by

$$\xi = \left\langle \left(\hat{A}+i\lambda\hat{B}\right)\psi \,\middle|\, \left(\hat{A}+i\lambda\hat{B}\right)\psi \right\rangle$$

which is real and positive since the inner product provides the length of the vector defined by

$$\left|\left(\hat{A}+i\lambda\hat{B}\right)\psi\right\rangle = \left(\hat{A}+i\lambda\hat{B}\right)|\psi\rangle$$

Assume that λ is a real parameter. Now working with the number ξ and using the definition of adjoint, namely

$$\langle \hat{O}f|g\rangle = \langle f|\hat{O}^{+}g\rangle,$$

we find

$$\begin{aligned}\xi &= \left\langle \psi \,\middle|\, \left(\hat{A}+i\lambda\hat{B}\right)^{+}\left(\hat{A}+i\lambda\hat{B}\right)\psi \right\rangle = \langle\psi|\left(\hat{A}+i\lambda\hat{B}\right)^{+}\left(\hat{A}+i\lambda\hat{B}\right)|\psi\rangle \\ &= \langle\psi|\left(\hat{A}^{+}-i\lambda\hat{B}^{+}\right)\left(\hat{A}+i\lambda\hat{B}\right)|\psi\rangle \\ &= \langle\psi|\left(\hat{A}\ -i\lambda\hat{B}\ \right)\left(\hat{A}+i\lambda\hat{B}\right)|\psi\rangle\end{aligned}$$

where the last step uses the Hermiticity of the operators $\hat{A},\hat{B}$. Multiply the operator terms in the bracket expression and suppress the reference to the wave function (for convenience) to obtain

$$\xi = \langle \hat{A}^2 \rangle - \lambda \langle \hat{C} \rangle + \lambda^2 \langle \hat{B}^2 \rangle \geq 0$$

which must hold for all values of the parameter λ. The minimum value of the positive real number ξ is found by differentiating with respect to the parameter λ.

$$\frac{\partial \xi}{\partial \lambda} = 0 \rightarrow \lambda = \frac{\langle \hat{C} \rangle}{2\langle \hat{B}^2 \rangle}$$

So that the minimum value of the positive real number ξ is given by

$$\xi_{min} = \langle \hat{A}^2 \rangle - \frac{1}{4} \frac{\langle \hat{C} \rangle^2}{\langle \hat{B}^2 \rangle} \geq 0$$

Multiplying through by $\langle \hat{B}^2 \rangle$ to get

$$\langle \hat{A}^2 \rangle \langle \hat{B}^2 \rangle \geq \frac{1}{4} \langle \hat{C} \rangle^2 \qquad (2.11.1)$$

The commutator $[\hat{A}, \hat{B}] = i\hat{C}$ holds for the two Hermitian operators defined by

$$\hat{A} \rightarrow \hat{A} - \langle \hat{A} \rangle \quad \hat{B} \rightarrow \hat{B} - \langle \hat{B} \rangle$$

As a result, Equation 2.11.1 becomes

$$\left\langle \left(\hat{A} - \langle \hat{A} \rangle \right)^2 \right\rangle \left\langle \left(\hat{B} - \langle \hat{B} \rangle \right)^2 \right\rangle \geq \frac{1}{4} \langle \hat{C} \rangle^2$$

However, the terms in the angular brackets are related to the standard deviations σ_a, σ_b respectively. The proof of the theorem is finished by taking the square root of the previous expression

$$\sigma_a \sigma_b \geq \frac{1}{2} \left| \langle \hat{C} \rangle \right|$$

Notice that this Heisenberg uncertainty relation involves the absolute value of the expectation value of the operator C. By its definition, the operator C is Hermitian and its expectation value is real.

Section 2.12: Quantum Mechanical Representations

A representation (or picture) in quantum theory refers to the objects of the theory that contain the dynamics; that is, the mathematical objects (wave vectors or operators) that contain the time-parameter. The Schrodinger representation uses wave functions that depend on time. The evolution of the particle or system is described by the wave function. The *act* of observing a particular quantity does not depend on when the observation is made. Intuitively, this point of view seems natural since the universe changes and not our eyes. The Schrodinger representation is very similar to classical electromagnetic theory for the wave motion of the photon field. The wave function resides in Hilbert space defined by a time-independent set of basis vectors. The wave function moves around in the vector space so that its components along the basis vector axes changes with time.

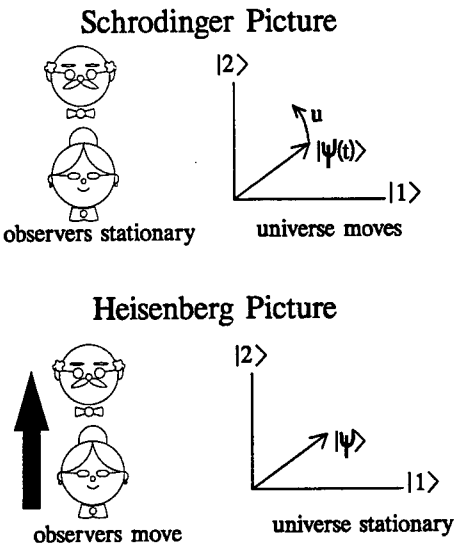


Figure 2.12.0: Intuitive picture for the Schrodinger and Heisenberg representations in quantum mechanics

The Heisenberg representation assigns all of the time dependence to the operators. This representation is somewhat similar to a classical particle theory. In classical mechanics, the dynamical variables such as momentum depend on time. For quantum theory, the Heisenberg wave functions do not depend on time. Intuitively, the wave vectors form the latticework of a stage that defines the allowed universe. The act of observing the "lattice work" depends on time. The wave function defines the extent of the physical system (i.e., it carries all the information on the properties of the particle). The Schrodinger representation allows the wave function to evolve in time which corresponds to the evolution (or motion) of the particle. The Heisenberg representation on the other hand places the dynamics in the operators.

The interaction representation assigns some time dependence to the operators and some to the wave functions. This representation is especially suited for an "open" system. A closed system can be defined as one for which the total energy (and the number of particles) contained within the system remains constant. For this case, imagine a system of atoms (maybe a chunk of semiconductor) with an external influence such as sun light. The motion of electrons in the system derives from the influence of internal forces and from the influence of the external source. The motion due to the internal sources is assumed trivial and the time dependence is assigned to the operators. The motion due to the external influence is assigned to the wavefunction. Without external sources, the wavefunctions are stationary similar to the Heisenberg representation. The demarcation between external and internal, in this example, is setup primarily for perturbation theory. The same concepts can be applied to a closed system so long as the Hamiltonian is divided into two parts.

Topic 2.12.1: The Schrodinger Representation

As previously discussed, the time-dependent wave function satisfies Schrodinger's equation.

$$\hat{H}|\psi(t)\rangle = i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle$$

or in a coordinate representation

$$\hat{H}\psi(x,t) = i\hbar\frac{\partial}{\partial t}\psi(x,t)$$

The wavefunction $|\psi(t)\rangle$ is expressed as an expansion in the basis of energy eigenfunctions. The basis functions are stationary (i.e., no time dependence); they are the stationary solutions of the time independent Schrodinger equation. Figure 2.12.1 shows the motion of the total wave function through Hilbert space. The wavefunction remains normalized and therefore the motion must be attributed to a unitary operator $\hat{u}$. The unitary operator, called the evolution operator, moves the initial wave function forward in time according to

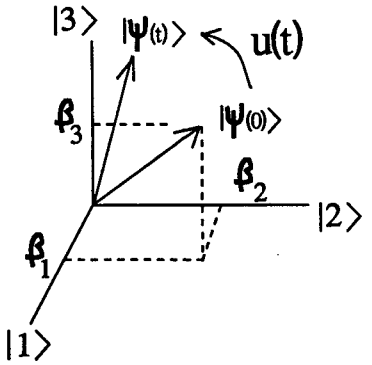


Figure 2.12.1: The wavefunction moves with respect to the energy eigenfunction basis set under the influence of the evolution operator.

$$\hat{u}(t,t_o)|\psi(t_o)\rangle = |\psi(t)\rangle$$

The evolution operator actually depends on the difference in time and is written as

$$\hat{u}(t,t_o) = \hat{u}(t-t_o)$$

Often the initial time is taken as zero $t_o = 0$

A formal expression for the evolution operator can be found by working with Schrodinger's equation. Assume that the Hamiltonian is independent of time (i.e., we have a close system). Schrodinger equation

$$\hat{H}|\psi(t)\rangle = i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle$$

can be formally integrated by writing it as

$$\frac{\partial}{\partial t}|\psi(t)\rangle = \frac{\hat{H}}{i\hbar}|\psi(t)\rangle$$

Consider the Hamiltonian operator to be similar to a constant and solve the simple differential equation to obtain

$$|\psi(t)\rangle = \exp\left(\frac{\hat{H}}{i\hbar}t\right)|\psi(0)\rangle = \hat{u}(t)|\psi(0)\rangle \qquad (2.12.1)$$

The operator

$$\hat{u}(t) = \exp\left(\frac{\hat{H}}{i\hbar}t\right)$$

is unitary (i.e., $\hat{u}^{-1} = \hat{u}^{+}$) since the Hamiltonian $\hat{H}$ is Hermitian. We know from the previous chapter that the exponential of an operator really stands for a Taylor expansion

using the operator as a variable. Equation 2.12.1 can also be written in the coordinate space as

$$\psi(x,t) = \exp\left(\frac{\hat{H}}{i\hbar}t\right)\psi(x,0) = \hat{u}(t)\psi(x,0)$$

The evolution operator plays a pivotal role for the Heisenberg representation.

Topic 2.12.2: Ehrenfest's Theorem for the Schrodinger Representation

In this topic, we discuss how an observed *value* (not the operator!) evolves in time for the Schrodinger representation. Ehrenfest's Theorem then shows how Schrodinger's quantum mechanics gives the usual results for classical mechanics. Basically the time dependence of a wave function must be related to the time dependence of observed values. The classical analog of a quantum mechanical dynamical variables involves an average over the quantum mechanical microscopic components. For example, the momentum of a classical point particle is the expectation value of the momentum operator. Newton's second law relates the time rate of change of the classical momentum to the applied force. Ehrenfest's theorem provides a method to recover Newton's second law by calculating the rate of change of the expectation value of the quantum mechanical momentum operator. The discussion by considering the rate of change of the expectation value of an operator using the Schrodinger picture. The derivation is carried all the way through; instead, the result is just stated.

Let $\hat{A} = \hat{A}(\vec{r},t)$ be an operator in the Schrodinger picture. Usually Shrodinger operators do not depend on time. One notable exception is the density operator. Suppose further that the wave vector $|\psi(t)\rangle$ is a solution to Schrodinger's equation (it can be a linear combination of the basis vectors as usual). The time rate of change of the expectation value of the operator is given by

$$\frac{d}{dt}\langle\hat{A}\rangle = \frac{d}{dt}\langle\psi|\hat{A}|\psi\rangle$$

where recall that inner products involve spatial integrals.

$$\frac{d}{dt}\langle\hat{A}\rangle = \frac{i}{\hbar}\langle[\hat{H},\hat{A}]\rangle + \left\langle\frac{\partial\hat{A}}{\partial t}\right\rangle \qquad (2.12.2)$$

Usually the expectation value of the time derivatives of the operator (last term) is zero for the Schrodinger picture.

Ehrenfest's theorem in Equation 2.12.2 can be intuitively understood. First consider Figure 2.12.2 For classical object, and applied force acts on the center of mass. An extended body such as a baseball consists of small individual masses denoted by

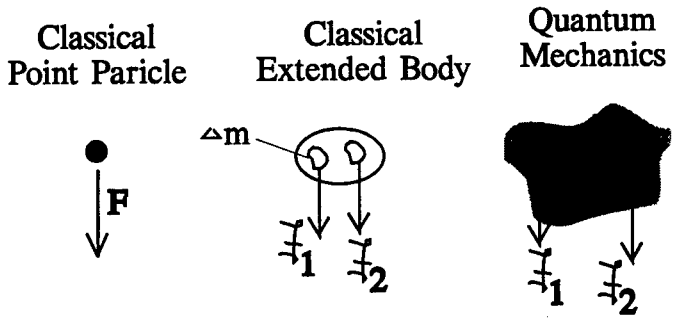


Figure 2.12.2: Left: force acting on a classical point mass. Middle: distributed force acting on a classical extended mass. Right: Force acting on a quantum mechanical object.

Δm_i. The total force $\vec{F}$ is made up of the individual forces acting on each mass. Define the quantity $\vec{F}$ to be the force per unit mass. The total force acting on the extended body is therefore given by

$$\vec{F} = \sum_i \vec{F}_i \Delta m_i$$

so that if a given region of the body has greater mass density then it might experience greater force. The figure shows a quantum mechanical object; the boundaries of the object are not very well-defined since the object is defined by a wave function. The object might be a single electron but it is pictured as spread out over a region of space. Once again define the total classical force as

$$\vec{F} = \sum_i \vec{F}_i \Delta m_i$$

but now the mass is not necessarily uniformly distributed across the region of space. The figure shows more mass near the center and less at the "boundaries." The amount of mass in a given region is proportional to the probability density of finding the electron in a small region. For the one dimensional case, write

$$\Delta m_i \sim p_r \, dx \sim \psi^* \psi \, \Delta x$$

therefore the total force is

$$\vec{F} \sim \sum_i \vec{F}_i \Delta m_i \sim \sum_i \psi^*(x_i) \vec{F}_i(x_i) \psi(x_i) \Delta x \sim \int \psi^*(x) \vec{F}(x) \psi(x) dx \sim \langle \vec{F} \rangle$$

Therefore, because the quantum mechanical object is necessarily spread across space, classical quantities like force and interaction energy do not occur at one specific point; instead they occur over the region of space. For this reason, the quantum mechanical operator is averaged over a region of space to produce the corresponding classical quantity. Furthermore, this shows that the time-dependence of the wave function translates to a time dependence of the corresponding classical quantity.

The figure at the right shows a wave packet traveling to the right with speed "v". The wavefunction clearly depends on time because it moves. The expectation value of the position operator $\hat{x}$ gives the position of the center of the wavepacket. Now because the wavepacket moves, the expectation value of the position operator must depend on time

$$\langle \hat{x} \rangle = \bar{x}(t)$$

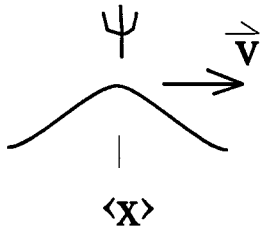


Figure 2.12.3: Example of an average that changes with time because the wavefunction depends on time.

Topic 2.12.3: The Heisenberg Representation

The Heisenberg representation assigns the dynamics to the operators. All of the wave functions, not just those in the basis set, are independent of time; none of the dynamics appears in the wave functions. The time-dependent operators are easy to find from those in the Schrodinger picture. We can proceed in two different ways. Perhaps the easiest method is to require all expectation values to remain unaffected by the particular picture. Suppose that the state of the system is represented by the ket $|\psi_s(t)\rangle$ in the Schrodinger picture (where "s" denotes Schrodinger). The expectation value of an operator $\hat{O}_s$ is given by

$$\langle\psi_s(t)|\hat{O}_s|\psi_s(t)\rangle = \langle\psi_h|\hat{u}^+\hat{O}_s\hat{u}|\psi_h\rangle$$

where the unitary operator $\hat{u}$ is the very same one used as the evolution operator for the Schrodinger picture. The origin of time is assumed to be t=0. The Heisenberg wave function is

$$|\psi_h\rangle = |\psi_s(0)\rangle$$

Recall that the unitary operator is given by

$$\hat{u}(t) = \exp\left(\frac{\hat{H}}{i\hbar}t\right)$$

The time-dependent Heisenberg operator is defined as

$$\hat{O}_h(t) = \hat{u}^+\hat{O}_s\hat{u}$$

Example: What is the Heisenberg representation of the Schrodinger Hamiltonian?
Solution: This is easy to find since the Schrodinger Hamiltonian commutes with itself

$$\hat{H}_h = \hat{u}^+\hat{H}_s\hat{u} = \exp\left(-\frac{\hat{H}_s}{i\hbar}t\right)\hat{H}_s\exp\left(\frac{\hat{H}_s}{i\hbar}t\right) = \hat{H}_s$$

So the Schrodinger and Heisenberg representations of the Hamiltonian are identical.

-- -- --

Next, the principal method of calculating the time evolution of the Heisenberg operators is discussed. The dynamics of the Heisenberg operators obtain through a commutation relation. It is this commutation relation that takes the place of the Schrodinger equation The commutation relation is reminiscent of the commutation relation found for the rate of change of the expectation value of an operator in the Schrodinger representation. The difference here is that it is not necessary to calculate the expectation value. In a sense, the operators in the Heisenberg representation are very similar to the dynamical variables in classical mechanics.

The rate of change of a Heisenberg operator is

$$\frac{d\hat{O}_h}{dt} = \frac{d}{dt}\left(\hat{u}^+\hat{O}_s\hat{u}\right) = \frac{d}{dt}\left\{\exp\left(-\frac{\hat{H}}{i\hbar}t\right)\hat{O}_s\exp\left(+\frac{\hat{H}}{i\hbar}t\right)\right\}$$

$$= \left\{\frac{d}{dt}\exp\left(-\frac{\hat{H}}{i\hbar}t\right)\right\}\hat{O}_s\exp\left(+\frac{\hat{H}}{i\hbar}t\right) + \exp\left(-\frac{\hat{H}}{i\hbar}t\right)\left\{\frac{\partial}{\partial t}\hat{O}_s\right\}\exp\left(+\frac{\hat{H}}{i\hbar}t\right)$$

$$+ \exp\left(-\frac{\hat{H}}{i\hbar}t\right)\hat{O}_s\left\{\frac{d}{dt}\exp\left(+\frac{\hat{H}}{i\hbar}t\right)\right\}$$

$$= \left\{-\frac{\hat{H}}{i\hbar}\exp\left(-\frac{\hat{H}}{i\hbar}t\right)\right\}\hat{O}_s\exp\left(+\frac{\hat{H}}{i\hbar}t\right) + \exp\left(-\frac{\hat{H}}{i\hbar}t\right)\left\{\frac{\partial}{\partial t}\hat{O}_s\right\}\exp\left(+\frac{\hat{H}}{i\hbar}t\right)$$

$$+ \exp\left(-\frac{\hat{H}}{i\hbar}t\right)\hat{O}_s\left\{\frac{\hat{H}}{i\hbar}\exp\left(+\frac{\hat{H}}{i\hbar}t\right)\right\}$$

$$= -\frac{\hat{H}}{i\hbar}\hat{O}_h + \hat{O}_h\frac{\hat{H}}{i\hbar} + \left(\frac{\partial}{\partial t}\hat{O}_s\right)_h$$

$$= \frac{i}{\hbar}\left[\hat{H},\hat{O}_h\right] + \left(\frac{\partial}{\partial t}\hat{O}_s\right)_h$$

where the time derivative of the Schrodinger operator is usually 0.

$$\frac{d\hat{O}_h}{dt} = \frac{i}{\hbar}\left[\hat{H},\hat{O}_h\right] + \left(\frac{\partial}{\partial t}\hat{O}_s\right)_h$$

Topic 2.12.4: The Interaction Representation

As previously mentioned, the interaction representation is a hybrid between the Schrodinger and Heisenberg representations. Both the operators and wave functions depend on time. This representation is especially used when a system interacts with an external force. The dynamics embedded in the wave function are due to the interaction. Therefore, the wave functions move in Hilbert space in response to the "extra" potentials imposed on the system. The operators carry the dynamics of the system that results when the interaction is turned-off.

Suppose the Hamiltonian for the system has the form

$$\hat{H} = \hat{H}_o + \hat{V}$$

where, for now, the Hamiltonian $\hat{H}_o$ is independent of time. Schrodinger's equation in operator form is

$$\hat{H}\left|\Psi_s(t)\right\rangle = i\hbar\frac{\partial}{\partial t}\left|\Psi_s(t)\right\rangle$$

or

$$\left(\hat{H}_o + \hat{V}\right)\left|\Psi_s(t)\right\rangle = i\hbar\frac{\partial}{\partial t}\left|\Psi_s(t)\right\rangle$$

Define the interaction wave vector

$$|\Psi_s(t)\rangle = \hat{u}|\Psi_I(t)\rangle$$

or equivalently

$$|\Psi_I(t)\rangle = \hat{u}^+|\Psi_s(t)\rangle$$

using the unitary operator previously defined for the Heisenberg representation

$$\hat{u}(t) = \exp\left(\frac{\hat{H}_o}{i\hbar}t\right)$$

The subscripts "s" and "I" stand for Schrodinger and Interaction, respectively. The Hamiltonian $\hat{H}_o$ defines the unitary evolution operator. Recall that the unitary operator $\hat{u}^+$ essentially removes the time dependence from the wave function $|\Psi_s(t)\rangle$ attributable to the Hamiltonian $\hat{H}_o$. However, there is still some time dependence left in the wave function due $\hat{V}$.

The Schrodinger equation can be written in the interaction representation. Starting with

$$\left(\hat{H}_o + \hat{V}\right)|\Psi_s(t)\rangle = i\hbar\frac{\partial}{\partial t}|\Psi_s(t)\rangle$$

substitute

$$|\Psi_s(t)\rangle = \hat{u}|\Psi_I(t)\rangle$$

to get

$$\left(\hat{H}_o + \hat{V}\right)\hat{u}(t)\,|\Psi_I(t)\rangle = i\hbar\frac{\partial}{\partial t}\hat{u}(t)\,|\Psi_I(t)\rangle \qquad (2.12.3)$$

Now, differentiate both terms on the right-hand side of Equation 2.12.3

$$\left(\hat{H}_o + \hat{V}\right)\hat{u}(t)\,|\Psi_I(t)\rangle = i\hbar\frac{\partial}{\partial t}\left[\exp\left(\frac{\hat{H}_o}{i\hbar}t\right)|\Psi_I(t)\rangle\right]$$

$$= i\hbar\left\{\frac{\hat{H}_o}{i\hbar}\exp\left(\frac{\hat{H}_o}{i\hbar}t\right)|\Psi_I(t)\rangle + \exp\left(\frac{\hat{H}_o}{i\hbar}t\right)\frac{\partial}{\partial t}|\Psi_I(t)\rangle\right\}$$

$$= \hat{H}_o\hat{u}\,|\Psi_I(t)\rangle + \hat{u}\frac{\partial}{\partial t}|\Psi_I(t)\rangle$$

We can cancel the terms involving $\hat{H}_o$ from both sides

$$\hat{V}\,\hat{u}(t)\,|\Psi_I(t)\rangle = i\hbar\,\hat{u}\frac{\partial}{\partial t}|\Psi_I(t)\rangle$$

Operating on both sides with the adjoint of the evolution operator

$$\hat{u}^+\hat{V}\,\hat{u}\,|\Psi_I(t)\rangle = i\hbar\frac{\partial}{\partial t}|\Psi_I(t)\rangle$$

or defining the interaction potential as

$$\hat{V}_I = \hat{u}^+\hat{V}\hat{u}$$

we can write

$$\hat{V}_I \left|\Psi_I(t)\right\rangle = i\hbar \frac{\partial}{\partial t} \left|\Psi_I(t)\right\rangle$$

As a result, the interaction wave function satisfies a Schrodinger-like equation with the interaction potential $\hat{V}_I$ taking the place of the Hamiltonian.

Section 2.13: The Harmonic Oscillator

The harmonic oscillator is familiar from classical mechanics. The oscillator consists of a particle capable of being displaced from an equilibrium position but the displacement initiates a linear restoring force. A linear restoring force is related to a quadratic potential V. The chapter focuses on the 1-D oscillator since the 3-D oscillator can be decomposed into three 1-D oscillators. Typically quantum theory focuses on the electron in a parabolic electrical potential. An anharmonic potential is not quadratic in the displacement and the harmonic oscillator is used as the starting point for the perturbation theory for approximation. For lasers, the main application for harmonic oscillators is not to the electron dynamics but instead, it is to light! The Hamiltonian for electromagnetic waves has a structure identical to that for the harmonic oscillator! The theory can be applied to a great many physical systems that have a second order differential equation. The solution to the Schrodinger equation proceeds using an algebraic approach inolving raising and lowering operators.

Topic 2.13.1: Introduction to the Classical and Quantum Harmonic Oscillators

Recall that the linear restoring force for a harmonic oscillator is

$$\vec{F} = \begin{cases} -\dfrac{\partial V}{\partial x}\hat{x} & 1-D \\ -\nabla V(\vec{r}) & 3-D \end{cases}$$

(x=0 is the equilibrium position of the particle) which is obtained from the quadratic potential

$$V = \begin{cases} \dfrac{1}{2}kx^2 & 1-D \\ \dfrac{1}{2}k\vec{r}^{\,2} & 3-D \end{cases}$$

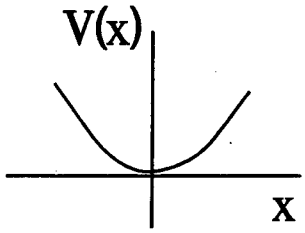


Figure 2.13.1: The quadratic potential for a harmonic oscillator

where the "spring constant" k>0. The classical Hamiltonian has the form

$$H_c = \frac{p^2}{2m} + \frac{1}{2}kx^2$$

where the dynamic variables x, p are considered to be independent of one another. Notice that Newton's second law can be recovered in the usual way though Hamilton's canonical equation

$$\dot{p} = -\frac{\partial H_c}{\partial x} = -kx = F$$

where F is the force applied by the potential. The Lagrangian shows that the momentum p is related to the velocity by $p = mv = m\dot{x}$ and the classical Hamiltonian is tranformed to

$$\frac{m}{2}\left(\frac{d\,x(t)}{dt}\right)^2 + \frac{1}{2}k\left(x(t)\right)^2 = E$$

where E is the total energy of the oscillator and x(t) is the position of the electron parameterized by the time t. The solution x(t) is well known from classical mechanics (assuming x(0)=0)

$$x(t) = A\sin(\omega_o t)$$

with $\omega_o^2 = \frac{k}{m}$. The angular frequency of oscillation is ω_o. Substituting into the second order equation provides

$$\frac{m}{2}A^2\omega^2\cos^2(\omega_o t) + \frac{1}{2}kA^2\sin^2(\omega_o t) = E$$

or, after substituting for ω_o^2,

$$kA^2 = 2E$$

or

$$A = \sqrt{\frac{2E}{k}} = \sqrt{\frac{2E}{m\omega_o^2}}$$

The amplitude "A" represents the point on the V(x) plot in Figure 2.13.2 where the kinetic energy is zero; i.e., where V(x) crosses the line of total energy E

$$E = \frac{1}{2}kx^2\Big|_{x=A} \rightarrow A = \sqrt{\frac{2E}{k}}$$

Classically, the particle can only be found in the range $x \in [-A, A]$ and never outside that region. The probability (density) ρ of finding the particle at a point x is also shown. Notice how the probability density is similar to a delta function near the endpoints of the motion; this behavior occurs because the particle slows down near the endpoints and spends more time there. Also notice that the angular frequency of oscillation ω is fixed by the parameters of the oscillator "k and m". Adjusting the amplitude A of oscillation changes the energy of the oscillator.

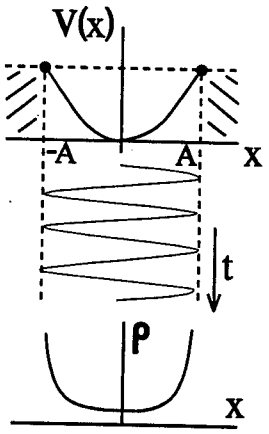


Figure 2.13.2: The sinusoidal motion of an electron in a harmonic potential. ρ is the probability density.

Several differences are expected for the quantum mechanical harmonic oscillator of mass m. First, the particle can be found in the classically forbidden region. The wavefunctions (shown in Figure 2.13.3) exponentially decay in these regions. Second, the probability density function decays to zero *near* the endpoints of motion and reaches its peak value near the center of the classical region [-A,A]. Third, for a particle in a quadratic potential, it is usually only possible to change the amplitude of oscillation unless an applied electric field (etc) can be used to change the quadratic potential. F*or EM waves*, the energy of an oscillator can be changed by changing the angular frequency (or wavelength) or by changing the amplitude (i.e., the number of quanta in the mode). For the EM harmonic oscillator in the next chapter, the "position x" and "momentum p" become the "in-phase" and "out-of-phase" electric fields. The wavefunctions in the EM case describe the probability of finding a particular value of the electric field.

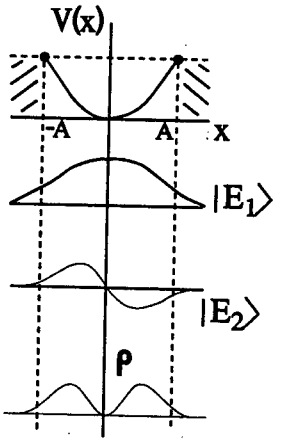


Figure 2.13.3: The quantum mechanical harmonic oscillator. The horizontal axis of all four plots is position x.

Figure 2.13.3 shows that the electron has a finite probability of being found in the classically forbidden region; classically, the electron doesn't have enough energy to enter the forbidden region. In principal, another harmonic oscillator could be moved close (but still in the forbidden region) to the first oscillator and the particle could escape its well and enter the second oscillator. This is an example of quantum mechanical tunneling. Semiconductor lasers have square wells that are close enough together to permit quantum tunneling between the wells.

Now here's another important difference between the classical and the quantum mechanical cases. For the classical case, the angular frequency is set by the parameters of the system

$$\omega_o = \sqrt{\frac{k}{m}}$$

However, for the QM oscillator, the angular frequency ω_n (corresponding to the energy levels) depends on the number of energy quanta in the oscillator. Figure 2.13.3 shows the frequency for the states $|E_1\rangle$ and $|E_2\rangle$ by looking at the "wavelength". Later discussion shows that the energy of the harmonic oscillator is quantized according to

$$E_n = \hbar\omega_o\left(n+\frac{1}{2}\right) \qquad n = 0,1,2,...$$

so that the angular frequency is

$$\omega_n = \omega_o\left(n+\frac{1}{2}\right) \qquad n = 0,1,2,...$$

Larger number of quanta "n" result in larger "displacements" from equilibrium (i.e., larger amplitude of oscillation) and larger regions where the electron can be found. Notice that ω_o appears in the equation for ω_n.

Topic 2.13.2: The Hamiltonian for the Quantum Harmonic Oscillator

The quantum mechanical Hamiltonian obtains from the classical one by replacing the dynamical variables x,p with the corresponding operators $\hat{x},\hat{p}$ (the caret "^" indicates an operator and not a unit vector). The Schrodinger equation

$$\hat{H}|\Psi(t)\rangle = i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle$$

is found by substituting the classical Hamiltonian

$$H_c = \frac{p^2}{2m}+\frac{1}{2}kx^2$$

with the operators replacing the classical dynamical variables

$$\left(\frac{\hat{p}^2}{2m}+\frac{1}{2}k\hat{x}^2\right)|\Psi(t)\rangle = i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle$$

Working in the x-coordinate representation (i.e., operate on the last equation with the "coordinate" projection operator "$\langle x|$"), we identify

$$x \leftrightarrow \hat{x} = x \qquad \text{and} \qquad p \leftrightarrow \hat{p} = \frac{\hbar}{i}\frac{\partial}{\partial x}$$

to obtain the coordinate representation of the Schrodinger equation

$$H\Psi(x,t) = i\hbar\frac{\partial\Psi(x,t)}{\partial t}$$

or

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi(x,t) + \frac{1}{2}kx^2\Psi(x,t) = i\hbar\frac{\partial}{\partial t}\Psi(x,t)$$

The boundary conditions for the harmonic oscillator Schrodinger wave equation consists of assuming that the wave function approaches zero as "x" goes to infinity

$$\Psi(x \to \pm\infty, t) \to 0$$

There are two methods for solving the harmonic oscillator Schrodinger wave equation. The first method uses a power series solution which becomes algebraically very involved. The second method uses raising and lowering operators and highlights the linear algebra. This book primarily uses the method of raising and lowering operators (commonly referred to as the algebraic approach).

Side comment:

As a side note, it is possible to mathematically show how the Shrodinger equation obtains from

$$\left(\frac{\hat{p}^2}{2m} + \frac{1}{2}k\hat{x}^2\right)|\Psi(t)\rangle = i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle$$

First operate on both sides using the x-coordinate projection operator $\langle x|$ to get

$$\langle x|\left(\frac{\hat{p}^2}{2m} + \frac{1}{2}k\hat{x}^2\right)|\Psi(t)\rangle = i\hbar\frac{\partial}{\partial t}\langle x|\Psi(t)\rangle$$

where the x-coordinate operator was moved past the time derivative. On the left-hand side, insert the unit operator defined by

$$1 = \int|x'\rangle dx'\langle x'|$$

between the Hamiltonian operator and the ket $|\Psi(t)\rangle$ to obtain

$$\langle x|\left(\frac{\hat{p}^2}{2m} + \frac{1}{2}k\hat{x}^2\right)\left(\int|x'\rangle dx'\langle x'|\right)|\Psi(t)\rangle = i\hbar\frac{\partial}{\partial t}\langle x|\Psi(t)\rangle$$

The x-terms can be moved under the integral since they do not depend on x'.

$$\int dx'\,\langle x|\left(\frac{\hat{p}^2}{2m} + \frac{1}{2}k\hat{x}^2\right)|x'\rangle\langle x'|\,|\Psi(t)\rangle = i\hbar\frac{\partial}{\partial t}\langle x|\Psi(t)\rangle \qquad (2.13.1)$$

Now assume that the momentum and position operators are diagonal in "x" so that

$$\langle x|\hat{p}^2|x'\rangle = \langle x|x'\rangle[\hat{p}(x')]^2 = \delta(x'-x)\left[\frac{\hbar}{i}\frac{\partial}{\partial x'}\right]^2 \quad \text{and} \quad \langle x|\hat{x}^2|x'\rangle = \delta(x'-x)[x']^2$$

since $\hat{x}|x'\rangle = x'|x'\rangle$. Therefore Equation 2.13.1 becomes

$$\int dx'\,\delta(x'-x)\left(\frac{-\hbar^2}{2m}\frac{\partial^2}{\partial x'^2}+\frac{1}{2}kx'^2\right)\langle x'|\Psi(t)\rangle = i\hbar\frac{\partial}{\partial t}\langle x|\Psi(t)\rangle$$

Integrating over the delta function yields

$$\left(\frac{-\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+\frac{1}{2}kx^2\right)\langle x|\Psi(t)\rangle = i\hbar\frac{\partial}{\partial t}\langle x|\Psi(t)\rangle$$

and, using $\langle x|\Psi(t)\rangle = \Psi(x,t)$, gives the desired results

$$\left(\frac{-\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+\frac{1}{2}kx^2\right)\Psi(x,t) = i\hbar\frac{\partial}{\partial t}\Psi(x,t)$$

Topic 2.13.3: Introduction to the Algebraic Solution to the Harmonic-Oscillator Schrodinger Equation for the Harmonic Oscillator

The algebraic approach to solving Schrodinger's equation for the harmonic oscillator is simpler than the power series approach. In addition, it provides a great deal of insight into the mathematical structure of the quantum theory. The algebraic approach uses "raising and lowering operators" (i.e., ladder operators, or sometimes called promotion and demotion operators). Ladder operators and creation/annihilation operators are very similar to each other and form the basis for most advanced studies in quantum theory.

The "raising a^+ and lowering operators a" map a Hilbert space into itself according to $a, a^+ : V \to V$. In the Schrodinger picture, these ladder operators are independent of time. Let $B_V = \{|0\rangle = |E_0\rangle, |1\rangle = |E_1\rangle, \ldots\}$ the basis set formed by the energy eigenvectors of the Hamiltonian. The notation used to define the basis set assumes non-degenerate eigenvalues E_n

$$\hat{H}|\phi_n\rangle = E_n|\phi_n\rangle$$

For every E_n, there is exactly one ϕ_n and one integer "n" (which typically starts at 0 for the harmonic oscillator). Therefore, there are three equivalent notations for the eigenstates

$$|n\rangle = |E_n\rangle = |\phi_n\rangle$$

An order is assumed to exist for the energy levels

$$E_0 < E_1 < E_2 < \cdots$$

The power series solution to the harmonic oscillator Schrodinger Equation shows that the energy eigenvalues are $E_n = \hbar\omega_o\left(n+\frac{1}{2}\right)$ where the angular frequency parameter is $\omega_o = \sqrt{\frac{k}{m}}$ and m,k are the mass of the particle and the spring constant, respectively. The wavefunctions incorporate the Hermite polynomials H_n. The first wave function, as an example, is

$$\langle x|E_0\rangle = \phi_0(x) = \left(\frac{\alpha}{\pi^{1/2}}\right)^{1/2} H_0(\alpha x)\exp\left(-\frac{\alpha^2 x^2}{2}\right) = \left(\frac{\alpha}{\pi^{1/2}}\right)^{1/2} \exp\left(-\frac{\alpha^2 x^2}{2}\right)$$

where

$$\alpha^2 = \left(\frac{m\omega_o}{\hbar}\right)$$

This topic assume the Hamiltonian describes a harmonic oscillator; however, it is possible to define operators similar to the ladder operators for other systems of interest (refer to Jayne-Cumming's model in Chapter 5).

Topic 2.13.4: Raising and Lowering Operators

The raising operator increases the basis vector by one according to the prescription

$$a^+|n\rangle = \sqrt{n+1}|n+1\rangle$$

while the lowering operator decreases the basis vector by one as given by

$$a|n\rangle = \sqrt{n}|n-1\rangle$$

The coefficients $\sqrt{n+1}, \sqrt{n}$ are the normalization constants (as will be seen later). When the lowering operator encounters the vacuum state (i.e., the lowest eigenvector which is $|0\rangle$) it gives 0 according to

$$a|0\rangle = 0$$

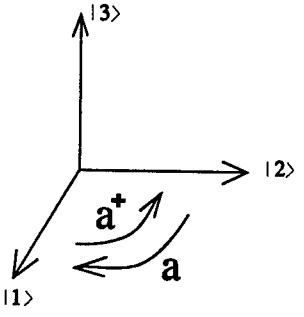


Figure 2.13.4: Raising and lowering operators move the harmonic oscillator from one state to another.

The vacuum state $|0\rangle$ corresponds to an electron state with no quanta of energy n=0. Interestingly, there is still energy in the vacuum state $E_0 = \hbar\omega_o/2$ which follows from the Hamiltonian eigenvalues given by $E_n = \hbar\omega_o\left(n+\frac{1}{2}\right)$. The vacuum energy corresponds to zero point motion of an oscillator and is the reason that absolute zero can never be achieved. It is also this vacuum energy that is responsible for the vacuum fluctuations of electric field, which produces spontaneous emission from an ensemble of excited atoms.

To be of importance, the raising and lowering operators must be related to the Hamiltonian for the harmonic oscillator. Typically these ladder operators are related to the position and momentum operators; all of the nice properties for the ladder operators follows from these definitions. The immediate goals are to (1) determine the Hamiltonian in terms of the ladder operators and (2) determine the basis functions using the ladder operators. Define the lowering and raising operators as

$$\hat{a} = \frac{m\omega_o}{\sqrt{2m\hbar\omega_o}}\hat{x} + \frac{i\hat{p}}{\sqrt{2m\hbar\omega_o}}$$

$$\hat{a}^+ = \frac{m\omega_o}{\sqrt{2m\hbar\omega_o}}\hat{x} - \frac{i\hat{p}}{\sqrt{2m\hbar\omega_o}}$$

The raising operator is found by taking the adjoint of the lowering operator and using the fact that both $\hat{x},\hat{p}$ are Hermitian since they correspond to observables. Notice that the raising and lowering operators are *not* Hermitian $\hat{a} \neq \hat{a}^+$. As a note sometimes people use $\hat{q}$ in place of $\hat{x}$. These two equations for the lowering and raising operators can be solved for the position and momentum operators to find

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega_o}}\left(\hat{a} + \hat{a}^+\right)$$

$$\hat{p} = -i\sqrt{\frac{m\omega_o\hbar}{2}}\left(\hat{a} - \hat{a}^+\right)$$

The commutation relations for the raising and lowering operators are necessary to complete the quantum prescription. These relations are easy to find based on those between the momentum and position operators

$$[\hat{x},\hat{x}] = 0 = [\hat{p},\hat{p}] \qquad [\hat{x},\hat{p}] = i\hbar$$

(1) The raising operator commutes with itself as does the lowering operator.

$$[\hat{a},\hat{a}] = 0 = [\hat{a}^+,\hat{a}^+]$$

(2) The raising operator does not commute with the lowering operator.

$$\begin{aligned}
[\hat{a},\hat{a}^+] &= \left[\frac{m\omega_o}{\sqrt{2m\hbar\omega_o}}\hat{x} + \frac{i\hat{p}}{\sqrt{2m\hbar\omega_o}}, \frac{m\omega_o}{\sqrt{2m\hbar\omega_o}}\hat{x} - \frac{i\hat{p}}{\sqrt{2m\hbar\omega_o}}\right] \\
&= \left(\frac{m\omega_o}{\sqrt{2m\hbar\omega_o}}\right)^2[\hat{x},\hat{x}] + \frac{[\hat{p},\hat{p}]}{2m\hbar\omega_o} + \frac{i}{2\hbar}[\hat{p},\hat{x}] - \frac{i}{2\hbar}[\hat{x},\hat{p}] \\
&= 0 + 0 + \frac{i}{2\hbar}(-i\hbar) - \frac{i}{2\hbar}(i\hbar) \\
&= 1
\end{aligned}$$

Notice that care was taken not to accidentally reversed the order of the operators $\hat{x}, \hat{p}$ and in the commutators. The results are very convenient and simple to remember

$$[\hat{a},\hat{a}^+] = 1$$

We will see later that for an ensemble of independent harmonic oscillators (assume that there are N of them), each has its own degrees of freedom $\hat{x}_i,\hat{p}_i$ that obey their own commutation relations.

$$[\hat{x}_i,\hat{x}_j] = 0 = [\hat{p}_i,\hat{p}_j] \qquad [\hat{x}_i,\hat{p}_j] = i\hbar\delta_{ij}$$

As a result, there will be raising and lowering operators for each oscillator

$$[\hat{a}_i,\hat{a}_j] = 0 = [\hat{a}_i^+,\hat{a}_j^+] \qquad [\hat{a}_i,\hat{a}_j^+] = \delta_{ij}$$

Topic 2.13.5: The Hamiltonian Using Ladder Operators

As previously mentioned, the raising and lowering operators map a basis vector into the next higher or lower one, respectively. *If the basis vectors are eigenvectors of the Hamiltonian, which is certainly true for our case, then the raising operator adds one*

quantum of energy while the lowering operator removes one quantum of energy. For example, Figure 2.13.5 shows the effect of the raising operator on an electron in a square well and in an atom.

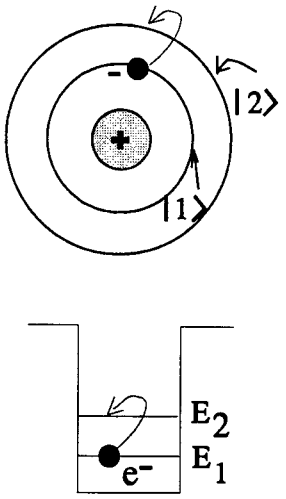


Figure 2.13.5: Physical examples showing the effect of a raising operator on an atom (top) and square well (bottom)

Using the definitions of the position and momentum operators, the Hamiltonian for the single harmonic oscillator can be written as

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega_o^2\hat{x}^2 = \frac{1}{2m}\left[-i\sqrt{\frac{m\omega_o\hbar}{2}}\left(\hat{a}-\hat{a}^+\right)\right]^2 + \frac{1}{2}m\omega_o^2\left[\sqrt{\frac{\hbar}{2m\omega_o}}\left(\hat{a}+\hat{a}^+\right)\right]^2$$

$$= -\frac{\hbar\omega_o}{4}\left(\hat{a}-\hat{a}^+\right)^2 + \frac{\hbar\omega_o}{4}\left(\hat{a}+\hat{a}^+\right)^2$$

$$= \frac{\hbar\omega_o}{4}\left\{-\hat{a}^2 + \hat{a}\hat{a}^+ + \hat{a}^+\hat{a} - \hat{a}^{+2} + \hat{a}^2 + \hat{a}\hat{a}^+ + \hat{a}^+\hat{a} + \hat{a}^{+2}\right\}$$

$$= \frac{\hbar\omega_o}{2}\left\{\hat{a}\hat{a}^+ + \hat{a}^+\hat{a}\right\} = \frac{\hbar\omega_o}{2}\left\{2\hat{a}^+\hat{a}+1\right\}$$

Notice in the third line that care was taken not to accidentally reversed the order of a,a^+ in any of the terms. The order of operators can only be changed by using the commutation relations. The last line was obtained by using

$$\left[\hat{a},\hat{a}^+\right] = 1 \quad \rightarrow \quad \hat{a}\hat{a}^+ = 1 + \hat{a}^+\hat{a}$$

As a result, the Hamiltonian for the single harmonic oscillator is written as

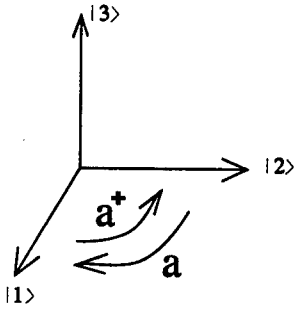
$$\hat{H} = \hbar\omega_o\left(\hat{a}^+\hat{a} + \frac{1}{2}\right) \qquad (2.13.2)$$

Topic 2.13.6: The Factors $\sqrt{n}$ and $\sqrt{n+1}$

This topic shows how the coefficients arise in

$$a^+|n\rangle = \sqrt{n+1}\,|n+1\rangle$$

$$a\,|n\rangle = \sqrt{n}\,|n-1\rangle$$



First, recall that the basis vectors and energy eigenvalues are defined so that

$$\hat{H}\,|n\rangle = E_n|n\rangle = \hbar\omega_o\left(n+\frac{1}{2}\right)|n\rangle$$

Compare this equation with the results of Equation 2.13.2 to see that when the quantity $\hat{a}^+\hat{a}$ operates on the state $|n\rangle$ we get

$$\hat{a}^+\hat{a}\,|n\rangle = n\,|n\rangle$$

where "n" is the number of quanta in the state (it is a real number). A number operator is therefore defined by

$$\hat{N} = \hat{a}^+\hat{a}$$

The number operator N gives the number of quanta in the harmonic oscillator state. The basis vectors $|0\rangle,|1\rangle,...$ have zero, one,... quanta of energy according to the formula $E_n = \hbar\omega_o\left(n+\frac{1}{2}\right)$.

The coefficients C_n and D_n in

$$a^+|n\rangle = C_n|n+1\rangle$$

$$a|n\rangle = D_n|n-1\rangle$$

can be found. To begin, first consider the string of equalities

$$D_n D_n^*\langle n-1|n-1\rangle = \left[D_n|n-1\rangle\right]^+\left[D_n|n-1\rangle\right] = \left[\hat{a}|n\rangle\right]^+\left[\hat{a}|n\rangle\right]$$

$$= \langle n|a^+a|n\rangle = \langle n|\hat{N}|n\rangle = \langle n|n|n\rangle = n\langle n|n\rangle$$

Now use the fact that all eigenvectors are normalized to one so that

$$\langle n-1|n-1\rangle = 1 = \langle n|n\rangle$$

Therefore, the coefficient D_n is

$$|D_n|^2 = n \quad \rightarrow D_n = \sqrt{n}$$

where a phase factor has been ignored. Similarly, an expression for C_n can be developed

$$C_n C_n^*\langle n+1|n+1\rangle = \left[C_n|n+1\rangle\right]^+\left[C_n|n+1\rangle\right] = \left[\hat{a}^+|n\rangle\right]^+\left[\hat{a}^+|n\rangle\right]$$

$$= \langle n|a\,a^+|n\rangle = \langle n|\ a^+a+1|n\rangle = \langle n|\hat{N}+1|n\rangle = \langle n|n+1|n\rangle = (n+1)\langle n|n\rangle$$

where a commutator has been used in the fifth term. Once again using the eigenvector normalization conditions and comparing both sides of the last equation

$$|C_n|^2 = n+1 \quad \rightarrow C_n = \sqrt{n+1}$$

as expected.

Some Commutation Relations

(1) $\left[\hat{H},a\right] = \hbar\omega_o\left[\hat{a}^+\hat{a},\hat{a}\right] = \hbar\omega_o\hat{a}^+\left[\hat{a},\hat{a}\right] + \hbar\omega_o\left[\hat{a}^+,\hat{a}\right]\hat{a} = -\hbar\omega_o\hat{a}$

(2) $\left[\hat{H},\hat{a}^+\right] = \hbar\omega_o\hat{a}^+\left[\hat{a},\hat{a}^+\right] = \hbar\omega_o\hat{a}^+$

(3) $\left[\hat{N},\hat{a}\right] = -\hat{a} \qquad \left[\hat{N},\hat{a}^+\right] = \hat{a}^+$

Topic 2.13.7: Energy Eigenvectors in Terms of the Raising Operator

The next topic shows how to obtain all of the energy eigenfunctions using the ladder operators; this is the simplest approach to solving the partial differential equation without using a series expansion. The procedure requires an expression for the energy eigenfunction in terms of the ladder operators and the vacuum state $|0\rangle$. The expression is obtained by repeated use of the relation $a^+|n\rangle = \sqrt{n+1}|n+1\rangle$.

$$|1\rangle = \frac{\hat{a}^+}{\sqrt{1}}|0\rangle$$
$$|2\rangle = \frac{\hat{a}^+}{\sqrt{2}}|1\rangle = \frac{(\hat{a}^+)^2}{\sqrt{2}\sqrt{1}}|0\rangle$$
$$\vdots$$
$$|n\rangle = \frac{(\hat{a}^+)^n}{\sqrt{n!}}|0\rangle$$

Topic 2.13.8: The Energy Eigenfunctions

There is a simple method for finding the energy eigenfunctions for the harmonic oscillator using the ladder operators. Starting with

$$0 = \hat{a}|0\rangle$$

operate on both sides using the bra operator $\langle x|$ and insert the definition for the lowering operator

$$0 = \langle x|a|0\rangle = \langle x|\frac{m\omega_o\hat{x}}{\sqrt{2\hbar m\omega_o}} + \frac{i\hat{p}}{\sqrt{2\hbar m\omega_o}}|0\rangle = \langle x|\frac{m\omega_o\hat{x}}{\sqrt{2\hbar m\omega_o}}|0\rangle + \langle x|\frac{i\hat{p}}{\sqrt{2\hbar m\omega_o}}|0\rangle \quad (2.13.3)$$

Factor out the constants from the brackets and use the relations

$$\langle x|\hat{x}|0\rangle = x\langle x|0\rangle = x\phi_0(x) \qquad \langle x|\hat{p}|0\rangle = \frac{\hbar}{i}\frac{\partial}{\partial x}\langle x|0\rangle = \frac{\hbar}{i}\frac{\partial}{\partial x}\phi_0(x)$$

where $\langle x|0\rangle = \phi_0(x)$ is the first energy eigenfunction in the set of eigenfunctions given by

$$\{\phi_0(x), \phi_1(x), \ldots\}$$

Equation 2.13.3 now provides

$$0 = \langle x|a|0\rangle = \frac{m\omega_o x}{\sqrt{2\hbar m\omega_o}}\phi_o(x) + \frac{\hbar}{\sqrt{2\hbar m\omega_o}}\frac{\partial}{\partial x}\phi_o(x)$$

which is a simple first-order differential equation

$$\frac{d\phi_0}{dx} + \frac{m\omega_o}{\hbar}x\phi_o = 0$$

The solution is easily seen to be

$$\phi_0(x) = \phi_0(0)\exp\left(-\frac{m\omega_o}{2\hbar}x^2\right)$$

which is the first energy eigenfunction. The normalization constant $\phi_0(0)$ is found by requiring the wave function to have unit length

$$1 = \langle\phi_0(x)|\phi_0(x)\rangle$$

which gives

$$\phi_0(x) = \left(\frac{m\omega_o}{\pi\hbar}\right)^{1/4}\exp\left(-\frac{m\omega_o}{2\hbar}x^2\right)$$

Now the other eigenfunctions obtain from $\phi_1(x)$ using the raising operator.

$$\phi_1(x)=\langle x|1\rangle=\langle x|\frac{a^+}{\sqrt{1}}|0\rangle=\langle x|\frac{m\omega_o\hat{x}}{\sqrt{2\hbar m\omega_o}}-\frac{i\hat{p}}{\sqrt{2\hbar m\omega_o}}|0\rangle$$

where the constants can be factored out and the coordinate representation can be substituted for the operators to get

$$\phi_1(x)=\frac{m\omega_o x}{\sqrt{2\hbar m\omega_o}}\langle x|0\rangle-\frac{\hbar}{\sqrt{2\hbar m\omega_o}}\frac{\partial}{\partial x}\langle x|0\rangle$$

$$=\frac{m\omega_o x}{\sqrt{2\hbar m\omega_o}}\phi_0(x)-\frac{\hbar}{\sqrt{2\hbar m\omega_o}}\frac{\partial}{\partial x}\phi_0(x)$$

There is no need to solve a differential equation to find this eigenfunction or the subsequent ones in the basis set. Differentiating $\phi_0(x)$ provides

$$\frac{\partial\phi_o}{\partial x}=\frac{\partial}{\partial x}\left(\frac{m\omega_o}{\pi\hbar}\right)^{1/4}\exp\left(-\frac{m\omega_o}{2\hbar}x^2\right)=-\frac{m\omega_o x}{\hbar}\phi_o(x)$$

Consequently the n=1 energy eigenfunction becomes

$$\phi_1(x)=\frac{m\omega_o x}{\sqrt{2\hbar m\omega_o}}\phi_0(x)+\frac{\hbar}{\sqrt{2\hbar m\omega_o}}\frac{m\omega_o x}{\hbar}\phi_0(x)=\frac{2\sqrt{m\omega_o}}{\sqrt{2\hbar}}x\phi_0(x)$$

$$=\frac{2\sqrt{m\omega_o}}{\sqrt{2\hbar}}\left(\frac{m\omega_o}{\pi\hbar}\right)^{1/4}x\exp\left(-\frac{m\omega_o x^2}{2\hbar}\right)$$

The n=2 energy eigenfunctions are found repeating the procedure using

$$\phi_2(x)=\frac{\hat{a}^+}{\sqrt{2}}\phi_1(x)$$

Notice that the above procedure only requires the relation between the ladder operators and the momentum/position operators. At this point, the energy eigenvalues can be found using Equation 2.13.2.

Section 2.14: Time-Dependent Perturbation Theory

Time-dependent perturbation theory is of primary importance for the study of the semiconductor laser. Perturbation theory provides calculation tools for approximating complicated Hamiltonians and solutions in terms of known solutions. Feynman diagrams are the result of applying perturbation calculations to quantum field theory especially as applied to modern elementary particles. This section presents several simple intuitive pictures for the transition of an electron (or other particle) from one energy level to another by absorbing or emitting a quantum of energy.

Topic 2.14.1: Physical Concept

Imagine a particle as occupying a specific energy level at the initial time. For example, Figure 2.14.1 shows an electron in the first energy level of an infinitely deep well, which might model a two- or three-level man-made atom. If the atom absorbs a quantum of energy from the incident electromagnetic field, which is the perturbation in this case, the electron makes a transition to a higher energy level as shown. The electron might be pictured as a small circle initially sitting on the unit vector $|E_1\rangle$. Notice this process is basically an exchange of energy and so the energy eigenstates are the appropriate basis vectors. Absorbing a photon from the incident electromagnetic field causes the electron to move from the first energy basis vector to the second one as shown. *A vitally important point to notice* is that the energy eigenstates are known and fixed! This means that the perturbation does not changed the size or angle of the eigenfunction in Hilbert space. Such a situation appears completely contrary to the approach taken for time independent perturbation theory! For time *independent* perturbation theory, an applied electric field changes the *unperturbed* energy eigenfunctions $|u_n\rangle$ into the *perturbed* eigenfunctions $|v_n\rangle$.

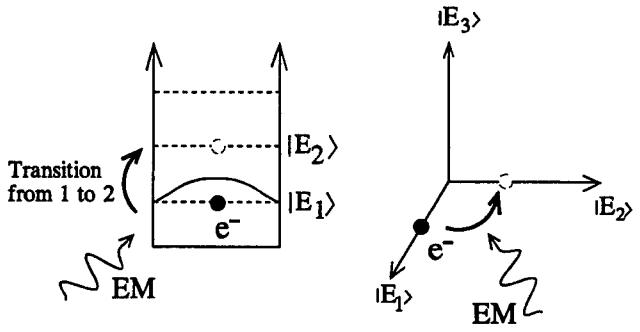


Figure 2.14.1: Electromagnetic waves induce transitions. The interaction energy is related to ladder-like operators.

Why are there differences between the two approaches to perturbation theory? For example, suppose that the depth of a quantum well is changed *very slowly* by slowly changing an applied electric field. In such a situation, time-*independent* perturbation theory should be applicable (maybe the change takes years) and yet time-*dependent* perturbation is explicitly used for cases like this one. It is possible to show the similarity between the two approaches (refer to the "Quantum Laser" by the present author).

The time-dependent perturbation theory can be presented in the Schrodinger, interaction or Heisenberg representation. The Hamiltonian for the system, which includes the perturbation $\hat{V}(x,t)$, is given by

$$\hat{H} = \hat{H}_o + \hat{V}(x,t)$$

and the unperturbed Hamiltonian $\hat{H}_o$ has the energy basis set $\{u_n = |n\rangle\}$ so that

$$\hat{H}_o|n\rangle = E_n|n\rangle$$

The Heisenberg representation is interesting because ladder-like operators acquire a time dependence; as a consequence, the transition from one state to another depends on time. However, this section first works with the Schrodinger representation where the wavefunction caries the full time-dependence. Consider a solution to Schrodinger's equation

$$|\psi(t)\rangle = \sum_n \beta_n(t)|n\rangle$$

where $|n\rangle = |u_n\rangle$ are the energy eigenfunctions of the unperturbed Hamiltonian $\hat{H}_o$. *Assume that the particle starts in state* $|i\rangle$ *at time t=0.* Recall that the components $\beta_n(t)$ essentially describe the probability of finding the particle in state $|n\rangle$ after a time t; actually,

$$\text{Prob}(n,t) = |\beta_n(t)|^2$$

So, if the particle starts in state $|i\rangle$ then $\beta_n(t)$ gives the probability of a transition from state $|i\rangle$ to state $|n\rangle$ after a time t.

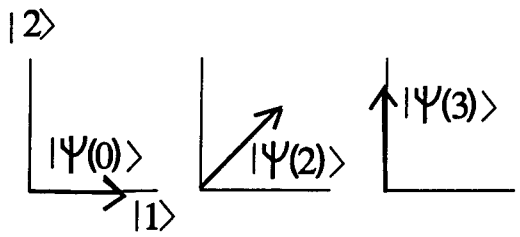


Figure 2.14.2: Evolution of the wavefunction under the action of the electromagnetic perturbation.

What does the wavefunction $|\psi(t)\rangle$ mean given the initial condition $|\psi(0)\rangle = |u_i\rangle$? Consider the sequence of plots in Figure 2.14.2. At time t=0, the wavefunction $|\psi(t)\rangle$ is totally along the $|1\rangle$ axis which says that the particle is definitely in the first energy eigenstate and so Prob(n=1,t=0) = 1. The second figure shows that at t=2, the electron is partly in both the first and second eigenstates. There is a probability of finding it in both. According to the figure,

$$\text{Prob}(n=1,t=2) = \text{Prob}(n=2,t=2) = 0.5$$

The third figure shows that at time t=3 that the electron is in state $|2\rangle$ alone since the wavefunction $|\psi(3)\rangle$ is parallel to the $|2\rangle$ axis . At t=3, the probability of finding the electron in state $|2\rangle$ is

$$\text{Prob}(n=2,t=3) = |\beta_2|^2 = 1$$

Notice that as the probability of finding the particle in state $|1\rangle$ decreases with time, that the probability of finding the particle in state $|2\rangle$ increases.

In the next topic, there are several key point to keep in mind. First, the basis vectors, which are the energy eigenfunctions of the unperturbed Hamiltonian H_o, are constant and do not change when the perturbing field is applied. Second, a first order

differential equation can be solved with an integrating factor (refer to Appendix 1 for a brief review). Third, the procedure solves for the components β_n of the wavefunction $|\psi(t)\rangle$ which are exact until the approximation $\beta(t) = \beta^{(0)}(t) + \beta^{(1)}(t) + \ldots$ is made. Forth, we the solution assumes that the particle starts in state $|i\rangle$ so that

$$\beta_n(0) = \beta_n^{(0)}(0) = \delta_{ni}$$

and

$$\beta_n^{(i)} = 0 \qquad i \geq 1$$

Fifth, the selection rules (i.e., the rules that determine the allowed transitions) are determined by the transition matrix elements

$$V_{fi} = \langle f|V|i\rangle$$

where i is the initial state and f is the final state.

The derivation is included for completeness sake. The uninterested reader should read through Equation 2.14.3 and then skip to the summary after Equation 2.14.10.

Topic 2.14.2: The Schrodinger Representation

In an effort to illustrate the concept of a particle making a transition from one fixed basis vector to another, we adopt a slightly cumbersome approach to the time-dependent perturbation formalism. However, the formulas will end up being the same as those in Yariv's book.

The Hamiltonian for the system, which includes the perturbation $\hat{V}(x,t)$, is given by

$$\hat{H} = \hat{H}_o + \hat{V}(x,t)$$

where Schrodinger's equation is

$$\hat{H}|\Psi(t)\rangle = i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle \quad \rightarrow \quad \left(\hat{H}_o + \hat{V}\right)|\Psi(t)\rangle = i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle \qquad (2.14.1)$$

the unperturbed Hamiltonian $\hat{H}_o$ has the energy basis set $\{u_n = |n\rangle\}$ so that

$$\hat{H}_o|n\rangle = E_n|n\rangle$$

We assume that $\{u_n = |n\rangle\}$ is also the basis set for the Hamiltonian $\hat{H}$. The reader familiar with boundary value problems and linear algebra will recognize this assumption as equivalent to treating the potential $\hat{V}(x,t)$ as a driving term for the partial differential equation. The Sturm-Liouville problem derives from the partial differential equation without the potential $\hat{V}(x,t)$ and the associated boundary conditions determine the basis set and the eigenvalues. It is essentially this step that relegates the perturbation to causing transitions between the basis vectors.

As usual, we write the solution to the full wave equation

$$\hat{H}|\Psi(t)\rangle = i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle \qquad (2.14.2)$$

as

$$|\Psi(t)\rangle = \sum_n \beta_n(t)|n\rangle \qquad (2.14.3)$$

Recall that the wave vector $|\Psi(t)\rangle$ moves in Hilbert space in response to the Hamiltonian $\hat{H}$ and that the components $\beta_n(t)$ are related to the probability of finding the electron in the state $|n\rangle$. As an *important point*, we assume that the particle is in state $|i\rangle$ at time t=0 ("i" in $|i\rangle$ is one of the integers 1,2,3...). We must substitute the expression 2.14.3 for the wave vector $|\Psi(t)\rangle$ into Equation 2.14.2 and solve for the components $\beta_n(t)$. Starting with

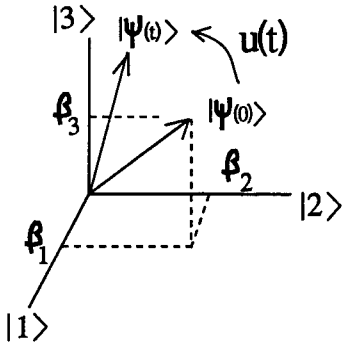


Figure 2.14.3: The motion of the wavefunction in Hilbert space.

$$\left(\hat{H}_o + \hat{V}\right)|\Psi(t)\rangle = i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle$$

we find

$$\left(\hat{H}_o + \hat{V}\right)\sum_n \beta_n(t)|n\rangle = i\hbar\frac{\partial}{\partial t}\sum_n \beta_n(t)|n\rangle$$

Move the unperturbed Hamiltonian and the potential inside the summation to find

$$\sum_n \beta_n(t)\left(E_n + \hat{V}\right)|n\rangle = i\hbar\sum_n \dot{\beta}_n(t)|n\rangle$$

where the dot over the symbol β indicates the time derivative. Operate on both sides of the equation with $\langle m|$ to find

$$\sum_n \beta_n(t)\left(E_n\langle m|n\rangle + \langle m|\hat{V}|n\rangle\right) = i\hbar\sum_n \dot{\beta}_n(t)\langle m|n\rangle$$

The orthonormality of the basis vectors

$$\langle m|n\rangle = \delta_{mn}$$

transforms the previous equation to

$$E_m\beta_m(t) + \sum_n \beta_n(t)\langle m|\hat{V}(x,t)|n\rangle = i\hbar\dot{\beta}_m(t)$$

which can be rewritten as

$$\dot{\beta}_m(t) - \frac{E_m}{i\hbar}\beta_m(t) = \frac{1}{i\hbar}\sum_n \beta_n(t)V_{mn}(t) \qquad (2.14.4)$$

where the matrix elements are

$$V_{mn}(t) = \langle m|\hat{V}(x,t)|n\rangle = \int dx\, u_m^* \hat{V}(x,t)u_n$$

for basis set consisting of functions of "x".

We must solve Equation 2.14.4 for the components $\beta_n(t)$ which is most easily handled by using an integrating factor $\mu_m(t)$. Rather than actually solve for the integrating factor, we will just state the results

$$\mu_m(t) = \exp\left(-\frac{E_m}{i\hbar}t\right) \qquad (2.14.5)$$

Multiplying the integrating factor on both sides of Equation 2.14.4, we can write

$$\mu_m\dot{\beta}_m - \frac{E_m}{i\hbar}\mu_m\beta_m = \frac{1}{i\hbar}\sum_n \mu_m\beta_n(t)V_{mn}$$

Noting that

$$\frac{d}{dt}(\mu_m\beta_m) = \dot{\mu}_m\beta_m + \mu_m\dot{\beta}_m \quad \text{and} \quad \dot{\mu}_m = -\frac{E_m}{i\hbar}\exp\left(-\frac{E_m t}{i\hbar}\right)$$

Equation 2.14.4 can finally be written as

$$\frac{d}{dt}[\mu_m(t)\beta_m(t)] = \mu_m(t)\dot{\beta}_m(t) - \frac{E_m}{i\hbar}\mu_m(t)\beta_m(t) = \frac{1}{i\hbar}\mu_m(t)\sum_n \beta_n(t)V_{mn}(t)$$

The reason for this maneuver will become apparent when the approximation is made. We need to solve the following equation for the components $\beta_n(t)$

$$\frac{d}{dt}[\mu_m(t)\beta_m(t)] = \frac{1}{i\hbar}\mu_m(t)\sum_n \beta_n(t)V_{mn}(t)$$

which obtains from the first and last terms of the previous equation. Assume that the perturbation starts at t=0 and integrate both sides with respect to time.

$$\mu_m(t)\beta_m(t) = \mu_m(0)\beta_m(0) + \frac{1}{i\hbar}\int_0 d\tau\,\mu_m(\tau)\sum_n \beta_n(\tau)V_{mn}(\tau)$$

Substituting for $\mu_m(t)$ and using the fact that the particle starts in state $|i\rangle$

$$\beta_n(0) = \delta_{ni}$$

we find

$$\beta_m(t) = \mu_m^{-1}(t)\delta_{mi} + \frac{\mu_m^{-1}(t)}{i\hbar}\sum_n \int_0 d\tau\,\mu_m(\tau)\beta_n(\tau)V_{mn}(\tau) \qquad (2.14.6)$$

Now we make the approximation by writing the components $\beta_n(t)$ as a summation

$$\beta_n(t) = \beta_n^{(0)}(t) + \beta_n^{(1)}(t) + \ldots$$

where the superscripts provide the order of the approximation in exactly the same manner as done for Equation 6.49.1 on page 49 in this chapter. Substituting the approximation for the components $\beta_n(t)$ into Equation 2.14.6 provides

$$\beta_m^{(0)}(t) + \beta_m^{(1)}(t) + \ldots = \mu_m^{-1}(t)\delta_{mi} + \frac{\mu_m^{-1}(t)}{i\hbar}\sum_n \int_0 d\tau\,\mu_m(\tau)[\beta_n^{(0)}(\tau) + \beta_n^{(1)}(\tau) + \ldots]V_{mn}(\tau)$$

Note that the terms $\beta_n^{(0)}V_{mn}$ are order "(1)" in the approximation since the interaction potential V_{mn} is already considered small. Equate corresponding orders of approximation to find

$$\beta_m^{(0)}(t) = \mu_m^{-1}(t)\delta_{mi} \qquad (2.14.7)$$

and

$$\beta_m^{(1)}(t) = \frac{\mu_m^{-1}(t)}{i\hbar}\sum_n \int_0 d\tau\,\mu_m(\tau)\beta_n^{(0)}(\tau)V_{mn}(\tau) \qquad (2.14.8)$$

and so on. Notice how Equation 2.14.8 invokes Equation 2.14.7 in the integral. So once we solve for the zeroth order approximation for the component, we can immediately find the first-order approximation. Higher order terms of work the same way.

Consider first the case for m=i which essentially suggest that the particle ends in the same state from which it started. Of course we are talking about the components $\beta_m(t)$ which are related to probability. The lowest order approximation gives (using Equation 2.14.5)

$$\beta_i^{(0)}(t)=\mu_i^{-1}(t)=\exp\left(\frac{E_i}{i\hbar}t\right) \tag{2.14.9}$$

Substituting Equation 2.14.9 into Equation 2.14.8 with m=i, we find

$$\begin{aligned}\beta_i^{(1)}(t)&=\frac{\mu_i^{-1}(t)}{i\hbar}\sum_n\int_0 d\tau\,\mu_i(\tau)\beta_n^{(0)}(\tau)V_{in}(\tau)\\&=\frac{\mu_i^{-1}(t)}{i\hbar}\sum_n\int_0 d\tau\,\mu_i(\tau)\delta_{in}\exp\left(\frac{E_i}{i\hbar}\tau\right)V_{in}(\tau)\\&=\frac{\mu_i^{-1}(t)}{i\hbar}\int_0 d\tau\,\mu_i(\tau)\exp\left(\frac{E_i}{i\hbar}\tau\right)V_{ii}(\tau)\end{aligned}$$

Substituting Equation 2.14.5 for the remaining integrating factors in the previous equation we find

$$\beta_i^{(1)}(t)=\frac{1}{i\hbar}\exp\left(\frac{E_i}{i\hbar}t\right)\int_0 d\tau\exp\left(+\frac{E_i}{i\hbar}\tau\right)\exp\left(\frac{E_i}{i\hbar}\tau\right)V_{ii}(\tau)=\frac{1}{i\hbar}\exp\left(\frac{E_i}{i\hbar}t\right)\int_0 d\tau\;V_{ii}(\tau)$$

So therefore the approximate value for $\beta_i(t)$ is

$$\beta_i(t)=\beta_i^{(0)}(t)+\beta_i^{(1)}(t)+\ldots=\exp\left(\frac{E_i}{i\hbar}t\right)+\frac{1}{i\hbar}\exp\left(\frac{E_i}{i\hbar}t\right)\int_0 d\tau\;V_{ii}(\tau)+\ldots$$

Next find the components $\beta_m(t)$ for a final state $|m\rangle$ that is not the same as the initial state $|i\rangle$; that is, m≠i.

$$\beta_m^{(0)}(t)=0$$

and

$$\begin{aligned}\beta_m^{(1)}(t)&=\frac{\mu_m^{-1}(t)}{i\hbar}\sum_n\int_0 d\tau\,\mu_m(\tau)\beta_n^{(0)}(\tau)V_{mn}(\tau)\\&=\frac{\mu_m^{-1}(t)}{i\hbar}\sum_n\int_0 d\tau\,\mu_m(\tau)\delta_{ni}\mu_i^{-1}(\tau)V_{mn}(\tau)\end{aligned}$$

Substitute Equation 2.14.5 for the integrating factors to get

$$\begin{aligned}\beta_m^{(1)}(t)&=\frac{1}{i\hbar}\exp\left(\frac{E_m}{i\hbar}t\right)\int_0 d\tau\exp\left(-\frac{E_m}{i\hbar}\tau\right)\exp\left(\frac{E_i}{i\hbar}\tau\right)V_{mi}(\tau)\\&=\frac{1}{i\hbar}\exp\left(\frac{E_m}{i\hbar}t\right)\int_0 d\tau\exp\left(-\frac{E_m-E_i}{i\hbar}\tau\right)V_{mi}(\tau)\end{aligned}$$

Many times differences such as E_m-E_i are written as

$$E_m\text{-}E_i=E_{mi}$$

and also

$$\omega_{mi}=\omega_m-\omega_i=\frac{E_m-E_i}{\hbar}=\frac{E_{mi}}{\hbar} \qquad (2.14.10)$$

The reader must keep track of the distinction between matrix elements and this new notation for differences between quantities. Using this notation

$$\beta_m^{(1)}(t)=\frac{1}{i\hbar}\exp\left(\frac{E_m}{i\hbar}t\right)\int_0 d\tau\,\exp\left(-\frac{E_{mi}}{i\hbar}\tau\right)V_{mi}(\tau)$$

Therefore, the components $\beta_m(t)$ for $m\neq i$ are approximately given by

$$\beta_m(t)=\beta_m^{(0)}(t)+\beta_m^{(1)}(t)+\ldots=0+\frac{1}{i\hbar}\exp\left(\frac{E_m}{i\hbar}t\right)\int_0 d\tau\,\exp\left(-\frac{E_{mi}}{i\hbar}\tau\right)V_{mi}(\tau)$$

In summary, the expansion coefficients in

$$|\Psi(t)\rangle=\sum_n \beta_n(t)|n\rangle$$

are given by

$$\beta_m(t)=\begin{cases}\exp\left(\frac{E_i}{i\hbar}t\right)+\frac{1}{i\hbar}\exp\left(\frac{E_i}{i\hbar}t\right)\int_0 d\tau\;V_{ii}(\tau)+\ldots & m=i\\ \frac{1}{i\hbar}\exp\left(\frac{E_m}{i\hbar}t\right)\int_0 d\tau\,\exp\left(-\frac{E_{mi}}{i\hbar}\tau\right)V_{mi}(\tau)+\ldots & m\neq i\end{cases}$$

Or, using shorthand notation of the type shown in Equation 2.14.10 for the angular frequency and energy of the states

$$\beta_m(t)=\begin{cases}e^{-i\omega_i t}+\frac{1}{i\hbar}e^{-i\omega_i t}\int_0 d\tau\;V_{ii}(\tau)+\ldots & m=i\\ \frac{1}{i\hbar}e^{-i\omega_m t}\int_0 d\tau\,e^{i\omega_{mi}\tau}\,V_{mi}(\tau)+\ldots & m\neq i\end{cases}$$

Topic 2.14.3: The Interaction Representation

The interaction representation for quantum mechanics is especially suited for situations where the Hamiltonian $\hat{H} = \hat{H}_o + \hat{V}(x,t)$ differs from the Hamiltonian $\hat{H}_o$ (with known eigenfunctions) by a small amount $\hat{V}(x,t)$. In a previous section, the operators and the wave functions are shown to depend on time in the interaction representation. The time-dependence induced by the Hamiltonian $\hat{H}_o$ is removed from the wave functions and placed in the operators. The wave functions retain the time-dependence induced by the potential $\hat{V}(x,t)$. Therefore, the motion of the wave function in Hilbert space reflects the dynamics embedded in the potential. If there is no potential $\hat{V}(x,t)$, then the wave functions are stationary in Hilbert space and the interaction representation reduces to the Heisenberg representation.

The Hamiltonian is

$$\hat{H} = \hat{H}_o + \hat{V}(x,t)$$

and the evolution operator is

$$\hat{u}(t) = \exp\left(\frac{\hat{H}_o}{i\hbar}t\right) \tag{2.14.11}$$

The interaction potential $\hat{V}_I$ in the interaction representation is

$$\hat{V}_I = \hat{u}^+ \hat{V} \hat{u}$$

The interaction wave function $|\Psi_I\rangle$ is

$$|\Psi_s\rangle = \hat{u}|\Psi_I\rangle$$

where $|\Psi_s\rangle$ is the usual Schrodinger wave function with the dynamics of the full Hamiltonian $\hat{H}$. The equation of motion for the interaction wave function is given by

$$\hat{V}_I\,|\Psi_I(t)\rangle = i\hbar\frac{\partial}{\partial t}|\Psi_I(t)\rangle \tag{2.14.12}$$

To find the wave function in the interaction representation, first formally integrate Equation 2.14.12 in the form

$$\frac{\partial}{\partial t}|\Psi_I(t)\rangle = \frac{1}{i\hbar}\hat{V}_I\,|\Psi_I(t)\rangle$$

to get

$$|\Psi_I(t)\rangle = |\Psi_I(0)\rangle + \frac{1}{i\hbar}\int_0^t d\tau\,\hat{V}_I(\tau)\,|\Psi_I(\tau)\rangle \tag{2.14.13}$$

where the interaction is assumed to start at t=0. The approximation consists of recognizing that the interaction wave function moves only a little in Hilbert space for a small interaction potential. *The 0^{th} order approximation* is

$$|\Psi_I(t)\rangle = |\Psi_I(0)\rangle$$

Subsequent approximations are obtained by substituting the wave functions into the integral. The *total first-order approximation* (including the 0^{th} order approximation) is

$$|\Psi_I(t)\rangle = |\Psi_I(0)\rangle + \frac{1}{i\hbar}\int_0 d\tau\, \hat{V}_I(\tau)\,|\Psi_I(0)\rangle \qquad (2.14.14)$$

The total second-order approximation is found by substituting Equation 2.14.14 into Equation 2.14.13 to obtain

$$|\Psi_I(t)\rangle = |\Psi_I(0)\rangle + \frac{1}{i\hbar}\int_0 d\tau\, \hat{V}_I(\tau)\left\{ |\Psi_I(0)\rangle + \frac{1}{i\hbar}\int_0 d\tau'\, \hat{V}_I(\tau')\,|\Psi_I(0)\rangle \right\}$$

which can be rearranged to give

$$|\Psi_I(t)\rangle = \left\{ 1 + \frac{1}{i\hbar}\int_0 d\tau\, \hat{V}_I(\tau) + \left(\frac{1}{i\hbar}\right)^2 \int_0 d\tau \int_0 d\tau'\, \hat{V}_I(\tau)\hat{V}_I(\tau') \right\} |\Psi_I(0)\rangle$$

Essentially, the unitary operator $\hat{u}^+(t)$ is being approximated by

$$\hat{u}^+ = \left\{ 1 + \frac{1}{i\hbar}\int_0 d\tau\, \hat{V}_I(\tau) + \left(\frac{1}{i\hbar}\right)^2 \int_0 d\tau \int_0 d\tau'\, \hat{V}_I(\tau)\hat{V}_I(\tau') \right\}$$

which is somewhat reminiscent of writing the operator as an exponential. For example, if the interaction potential were independent of time (but it is not) then the operator would reduce to

$$\hat{u}^+ \underset{x}{=} 1 + \frac{\hat{V}_I t}{i\hbar} + \left(\frac{\hat{V}_I t}{i\hbar}\right)^2 + \ldots = \exp\left(\frac{\hat{V}_I t}{i\hbar}\right)$$

Section 2.15: Translation Operators

Common mathematical operations such as rotating or translating coordinates are handle by operators in the quantum theory. Previous sections in this chapter show that states are transformed by the application of a single unitary operator whereas "operators" are transformed through a similarity transformation. The translation and rotation operators are unitary and have arguments that involve a Hermitian operator that is conjugate to the variable being translated or rotated (refer to Goldstein's book).

Topic 2.15.1: Exponential Form of the Translation Operator

Let $\hat{x}$ and $\hat{p}$ be the position operator and an operator defined in terms of a derivative

$$\hat{p} = \frac{1}{i}\frac{\partial}{\partial x}$$

which is the "position" representation of $\hat{p}$. The position representation of $\hat{x}$ is x. The operator $\hat{p}$ is Hermitian. The coordinate kets satisfy

$$\hat{x}|x\rangle = x|x\rangle$$

The operators satisfy commutation relations

$$[\hat{x}, \hat{p}] = [\hat{x}\hat{p} - \hat{p}\hat{x}] = i$$

as is easily verified by performing the calculation using a position representation

$$[\hat{x}, \hat{p}]f(x) = [\hat{x}\hat{p} - \hat{p}\hat{x}]f(x) = x\frac{1}{i}\frac{\partial}{\partial x}f - \frac{1}{i}\frac{\partial}{\partial x}(xf) = x\frac{1}{i}\frac{\partial}{\partial x}f - \frac{x}{i}\frac{\partial f}{\partial x} - \frac{1}{i}f = i\,f$$

comparing both sides, we see that the *operator* equation $[\hat{x}, \hat{p}] = i$ holds.

This topic shows that the exponential $e^{-ia\hat{p}}$ translates the coordinate system according to

$$\hat{T}(\eta)f(x) = e^{-i\eta\hat{p}}f(x) = f(x-\eta)$$

The proof starts by working with a small displacement ξ_k and calculating the Taylor expansion about the point "x"

$$f(x+\xi_k) \cong f(x) + \frac{\partial f(x)}{\partial x}\xi_k + \ldots = \left(1 + \xi_k\frac{\partial}{\partial x} + \ldots\right)f(x)$$

Substituting the operator for the derivative

$$\hat{p} = \frac{1}{i}\frac{\partial}{\partial x}$$

gives

$$f(x+\xi_k) = \left(1 + \xi_k\frac{\partial}{\partial x} + \ldots\right)f(x) = \left(1 + i\xi_k\hat{p} + \ldots\right)f(x)$$
$$= \exp(+i\xi_k\hat{p})f(x)$$

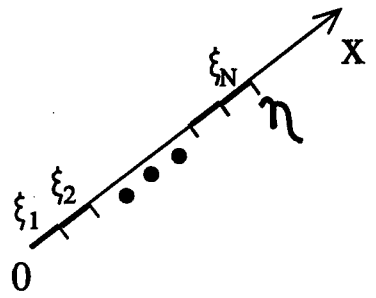


Figure 2.15.1: The displacement η is made of small infinitesimal displacements ξ_k.

Now, by repeated application of the infinitesimal translation operator, the entire displacement η can be built.

$$f(x+\eta)=\prod_k \exp\left(+i\xi_k \hat{p}\right) f(x)=\exp\left(\sum_k i\xi_k \hat{p}\right) f(x)=\exp\left(i\eta\hat{p}\right) f(x)$$

So the exponential with the momentum operator provides a translation. Replacing η with $-\eta$ provides

$$\hat{T}(\eta) f(x)=e^{-i\eta\hat{p}} f(x)=f(x-\eta)$$

Note that the translation operator is unitary $\hat{T}^{+}=\hat{T}^{-1}$ for η real since $\hat{p}$ is Hermitian. Also note that $\hat{T}^{+}(-\eta)=\hat{T}(\eta)$.

Topic 2.15.2: Translation of the Position Operator

This topic show that

$$\hat{T}^{+}(\eta)\hat{x}\,\hat{T}(\eta)=\hat{x}-\eta$$

where $\hat{T}(\eta)=e^{-i\eta\hat{p}}$. This is easy to show using the operator expansion theorem in Section 5.x.x which is

$$e^{\eta\hat{A}}\hat{B}e^{-\eta\hat{A}}=\hat{B}+\frac{\eta}{1!}\left[\hat{A},\hat{B}\right]+\frac{\eta^2}{2!}\left[\hat{A},\left[\hat{A},\hat{B}\right]\right]+\ldots$$

Using $\hat{A}=i\hat{p}$ and the commutation relations $[\hat{x},\hat{p}]=i$, we find

$$e^{i\eta\hat{p}}\hat{x}e^{-i\eta\hat{p}}=\hat{x}+\frac{\eta}{1!}\left[i\hat{p},\hat{x}\right]+\frac{\eta^2}{2!}\left[i\hat{p},\left[i\hat{p},\hat{x}\right]\right]+\ldots=\hat{x}-\eta$$

Topic 2.15.3: Translation of the Position-Coordinate Ket

The position-coordinate ket $|x\rangle$ is an eigenvector of the position operator $\hat{x}$

$$\hat{x}|x\rangle=x|x\rangle$$

What position-coordinate ket $|\phi\rangle$ is an eigenvector of the translated operator

$$\hat{T}^{+}(\eta)\hat{x}\,\hat{T}(\eta)=\hat{x}-\eta$$

that is, what is the state $|\phi\rangle=\hat{T}^{+}(\eta)|x\rangle$? The eigenvector equation for the translated operator $\hat{x}_T=\hat{T}^{+}\hat{x}\,\hat{T}$ is

$$\hat{x}_T|\phi\rangle=\hat{T}^{+}(\eta)\hat{x}\,\hat{T}(\eta)|\phi\rangle=\left[\hat{T}^{+}(\eta)\hat{x}\,\hat{T}(\eta)\right]\hat{T}^{+}(\eta)|x\rangle=\hat{T}^{+}(\eta)\,\hat{x}|x\rangle=x\,\hat{T}^{+}(\eta)|x\rangle=x|\phi\rangle$$

However, we know the translated operator is $\hat{x}_T=\hat{x}-\eta$ and therefore the previous equation provides

$$x|\phi\rangle=\hat{x}_T|\phi\rangle=(\hat{x}-\eta)|\phi\rangle=(\phi-\eta)|\phi\rangle$$

Comparing both sides, we see $\phi=x+\eta$ which therefore show that the translated position vector is

$$|\phi\rangle=\hat{T}^{+}(\eta)|x\rangle=|x+\eta\rangle$$

Topic 2.15.4: Example Using the Dirac Delta Function

Show that

$$|\phi\rangle = \hat{T}^{+}(\eta)|x'\rangle = |x'+\eta\rangle$$

using the results of Topic 2.16.3 and the fact that the position-ket represents the Dirac Delta function in Hilbert space

$$|x'\rangle \equiv |\delta(\bullet - x')\rangle$$

where "$\bullet$" represents the missing variable. If "x" is a coordinate on the x-axis then

$$\langle x|x'\rangle \equiv \int_{-\infty}^{\infty} d\varsigma\, \delta(\varsigma - x)\delta(\varsigma - x') = \delta(x - x')$$

Applying the translation operator in the x-representation using the results of Topic 2.15.3

$$\langle x|\hat{T}(\eta)|x'\rangle = e^{-i\eta\hat{p}_x}\langle x|x'\rangle = e^{-i\eta\hat{p}_x}\delta(x - x') = \delta(x - \eta - x') = \langle x|x'+\eta\rangle$$

Evidently

$$\hat{T}(\eta)|x'\rangle = |x'+\eta\rangle$$

Section 2.16: The Density Operator

The density operator and its associated equation of motion provided alternate formulation for a quantum mechanical system. The density operator combines both quantum mechanics and statistical mechanics into one mathematical object. The quantum mechanical part of the density operator is related to the usual quantum mechanical wave function, which accounts for the inherent uncertainty for a particle. The statistical mechanics portion describes the lack of knowledge of the exact wave function for the particle or system because of macroscopic statistical processes.

Topic 2.16.1: Introduction to the Density Operator

Previous topics in quantum mechanics pretend to know the initial wave function of a particle or system. For example, the exact decomposition of an initial wave function is

$$|\psi(0)\rangle = 0.9|u_1\rangle + 0.43|u_2\rangle$$

(the expansion coefficients are known exactly). The quantum mechanical probability of finding the electron in the first eigenstate is given by

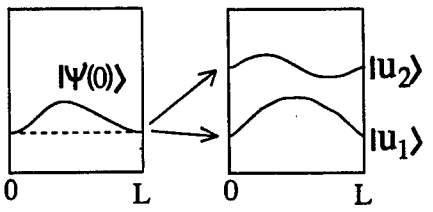


Figure 2.16.1: Spectral decomposition of the wavefunction

$$|\langle u_1|\psi(0)\rangle|^2 = (0.9)^2 = 81\%$$

Similarly, the quantum mechanical probability that the electron is in the second eigenstate is given by

$$|\langle u_2|\psi(0)\rangle|^2 = (0.43)^2 = 19\%$$

The these probabilities are known with certainty since the decomposition of the initial wave function $|\psi(0)\rangle$ and the coefficients 0.9 and 0.43 are known with 100% certainty. The wave function $|\psi\rangle$ is assumed to satisfy the time-dependent Schrodinger wave equation while the basis states satisfy the time-independent Schrodinger wave equation

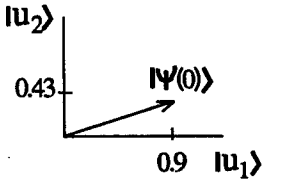


Figure 2.16.2: The expansion coefficients are known exactly.

$$\hat{H}|\psi\rangle = i\hbar\frac{\partial}{\partial t}|\psi\rangle \qquad \hat{H}|u_1\rangle = E_1|u_1\rangle \qquad \hat{H}|u_2\rangle = E_2|u_2\rangle$$

What if the initial preparation of the system is not exactly known? Suppose that an attempt is made to prepare a number of identical systems. Figure 2.16.3 shows four (out of an infinite number) of the systems, each of which has the same width L. The fact that the boundary conditions are the same for each implies that the basis sets are the same for each. The figure

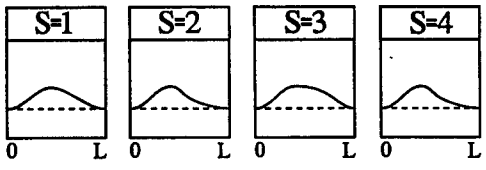


Figure 2.16.3: An ensemble of boxes containing one electron.

indicates that the attempt failed to produce identical wave functions $|\psi\rangle$. Let the wave function for system S be denoted by $|\psi_s\rangle$. Then the wave function $|\psi_s\rangle$ for each system has different coefficients, as for example,

$$|\psi_1\rangle = 0.98|u_1\rangle + 0.19|u_2\rangle$$
$$|\psi_2\rangle = 0.90|u_1\rangle + 0.43|u_2\rangle$$
$$|\psi_3\rangle = 0.95|u_1\rangle + 0.31|u_2\rangle$$
$$|\psi_4\rangle = 0.90|u_1\rangle + 0.43|u_1\rangle$$

Notice that system S=2 and system S=4 both have the same wave function which is also shown in Figures 2.16.3 and 2.16.4.

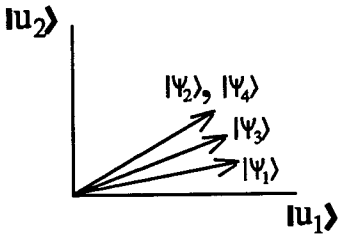


Figure 2.16.4: The wave functions for four of the systems in the ensemble plotted in Hilbert space.

What is the actual wave function $|\psi\rangle$ that describes the system? The answer is that there is no *actual* $|\psi\rangle$ -- only an average wavefunction. In fact, if many such systems had been prepared, then it is only appropriate to discuss the probability that the system has a certain wave function. For example, for the four systems described above, the probability of each type is given by

$$P(S=2)) = \frac{1}{2} \qquad P(S=1) = \frac{1}{4} \qquad P(S=3) = \frac{1}{4}$$

which is an assignment made to account for some type of unknown macroscopic influence. For example, maybe the electrons were injected into each system with varying degrees of energy which results in the wavefunction having slightly different spectra. *Notice that for convenience, the systems S=2 and S=4 are now called S=2 since the wave functions are identical.* Three vectors in Hilbert space now represent the four systems since the probability of each vector is also given.

It would be nice to have a single mathematical object that describes both the microscopic and macroscopic processes for the system. To start, define an average vector to represent the system

$$|\overline{\psi}\rangle = \text{average}|\psi\rangle = \sum_s P_s|\psi_s\rangle \qquad (2.16.1)$$

There is a similar definition for an average wavefunction with a continuously varying parameter. The definition makes use of the average for the quantity "x_i" or "x" written as

$$\langle x_i\rangle = \sum_i x_i P_i \qquad \langle x\rangle = \int dx\, x P(x)$$

for the discrete and continuous cases respectively. Equation 2.16.1 is a wave function for a system that represents the average system in the ensemble. *The wave functions* $|u_s\rangle$ *represent the quantum mechanical stochastic processes while the probabilities* P_S *represent the macroscopic (i.e., statistical mechanical) stochastic processes.*

Now here comes the unusual part. If the wave function is known exactly, then the probabilities can be calculated exactly using the quantum mechanical probability density

$$\psi^*(x)\psi(x)$$

which is obviously projected into coordinate space (it could have been projected into Fourier space or any other applicable space). The key here is that these probabilities can be calculated exactly (it's a little odd to be combining the words "exact" and "probability"). Now extend this idea of probability using the ensemble of systems. For convenience, *change notation* and let P_ψ be the probability of finding one of the systems to have a wave function of $|\psi\rangle$. In the old notation, P_s is the probability of $|\psi_S\rangle$; however, the "S" can be dropped so long as it is understood that the wave function $|\psi\rangle$ represents the wave function of one of the systems and not be average wave function as used previously. Define an average probability density function according to

$$\text{average}(\psi^*\psi) = \sum_\psi P_\psi \left(\psi^*(x)\psi(x)\right) \tag{2.16.2}$$

This formula contains both the quantum mechanical probability density $\psi^*\psi$ and the macroscopic probability P_ψ. Equation 2.16.2 also assumes that there is a discrete number of possible wave functions $|\psi_S\rangle$. It might be that there are so many different wave functions that the wave functions essentially form a continuum (i.e., "S" is a continuously varying parameter). In such a case, it is possible to use the classical probability density ρ_ψ, which describes probability per unit interval of finding a particular wave function.

$$\text{average}(\psi^*\psi) = \int dS\, \rho_S \left(\psi_S^*(x)\psi_S(x)\right)$$

The discussion continues with Equation 2.16.2 since it contains all of the essential ingredients.

Rearranging Equation 2.16.2 provides a "way to think of the average"

$$\text{average}(\psi^*\psi) = \sum_\psi P_\psi \psi^*(x)\psi(x) = \sum_\psi P_\psi\, \psi(x)\psi^*(x)$$

$$= \sum_\psi P_\psi \langle x|\psi\rangle\langle\psi|x\rangle$$

$$= \langle x| \left\{ \sum_\psi P_\psi |\psi\rangle\langle\psi| \right\} |x\rangle$$

We define the density operator to be

$$\hat{\rho} = \sum_\psi P_\psi |\psi\rangle\langle\psi|$$

The density operator represents

(1) the quantum mechanical probability through the wave functions

$$|\psi(t)\rangle = |\psi_S(t)\rangle$$

where $|\psi\rangle$ is related to the quantum mechanical probability which describes our lack of knowledge of microscopic quantities such as location of the electron.

(2) the statistical probability through $P_\psi = P_S$ where P_ψ represents an ensemble probability and describes our incomplete knowledge about macroscopic quantities and the system preparation.

As will be seen later, the matrix element

$$\rho_{aa} = \langle a|\rho|a\rangle = \langle u_a|\rho|u_a\rangle$$

where $\{|a\rangle = |u_a\rangle\}$ are the basis vectors (i.e., the energy eigenvectors), give the probability that electron can be found in the state $|a\rangle$. The off-diagonal elements are related to the polarization. The operator

$$\hat{\rho} = \sum_\psi P_\psi |\psi\rangle\langle\psi| \tag{2.16.3}$$

functions as a joint probability density.

As a note, it is possible that the initial wavefunction is exactly known (a pure state). Macroscopic stochastic processes can influence the initally pure wavefunction and end up with a state with a number of different P_ψ.

By the way, it is the definition of the density operator in Equation 2.16.3 that most closely resembles the P-distribution for quantum optics.

Topic 2.16.2: Basis Expansion of the Density Operator

The density operator can be written as sum over basis vectors and projectors. The density operator $\hat{\rho}$ has a range and domain within a single vector space. Suppose the vector space is spanned by the set of basis vectors $\{|m\rangle = u_m\}$. It is common to use the energy eigenfunctions as the basis set. Section 2.4 shows that the density operator can be written as

$$\hat{\rho} = \sum_{mn} \rho_{mn} |m\rangle\langle n| \tag{2.16.4}$$

where $\{\langle n|\}$ are the basis vectors for the dual space of projection operators (i.e. $\langle n| = |n\rangle^+$). Recall that ρ_{mn} are the matrix elements (that is, ρ_{mn} form the density matrix) since

$$\langle a|\hat{\rho}|b\rangle = \sum_{mn} \rho_{mn} \langle a|m\rangle\langle n|b\rangle = \sum_{mn} \rho_{mn} \delta_{am} \delta_{bn} = \rho_{ab}$$

where $|a\rangle, |b\rangle$ are basis vectors. This topic shows how the density operator is expanded in a basis set using two different approaches. The next topic shows several examples.

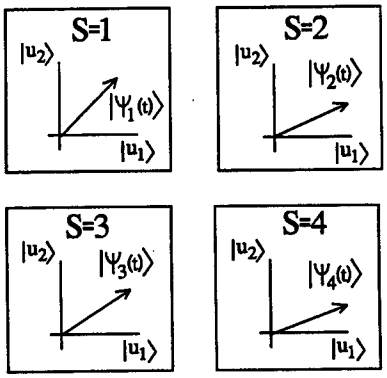


Figure 2.16.5: The example from the previous topic.

As previously mentioned, the density operator provides two types of average. The first is the quantum

mechanical average and the second is the ensemble average. For the ensemble average, imagine a large number of systems prepared as nearly the same as possible. Imagine a collection of wavefunctions $\{|\psi_S(t)\rangle\}$ with one for each different system S. Again, imagine that P_s is the probability of finding a particular wavefunction $|\psi_S(t)\rangle$. Assume that all of the systems can be described by vector spaces spanned by the set $\{|m\rangle = u_m\}$. Each wavefunction $|\psi_S(t)\rangle$ can be expanded in the complete orthonormal basis set for each system

$$|\psi_S(t)\rangle = \sum_m \beta_m^{(S)}(t)|m\rangle \qquad (2.16.5)$$

The superscript (S) on each expansion coefficient refers to a different system. However, there is one set of basis vectors that is applied to all of the systems S in the ensemble of systems. Therefore, the coefficients are not all the same $\beta_m^{(a)} \neq \beta_m^{(b)}$ in order that the wavefunction of system "a" is different from the wavefunction of system "b".

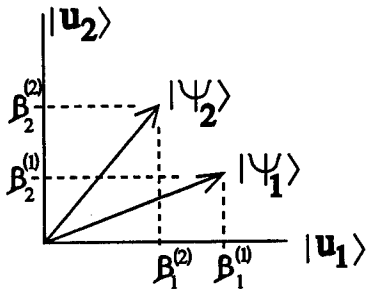


Figure 2.16.6: Two different wavefunctions in the ensemble. Each wavefunction has a different set of expansion coefficients.

The definition of the density operator provides

$$\hat{\rho}(t) = \sum_S P_S |\psi_S(t)\rangle\langle\psi_S(t)| \qquad (2.16.6)$$

Notice that the density operator in the Schrodinger representation can depend on time since the wave functions depend on time. The definition of adjoint gives

$$\langle\psi_S(t)| = |\psi_S(t)\rangle^+ = \left[\sum_n \beta_n^{(S)}|n\rangle\right]^+ = \sum_n \beta_n^{(S)*}\langle n| \qquad (2.16.7)$$

Substituting Equation 2.17.5 and 2.17.7 into Equation 2.17.6, we obtain

$$\hat{\rho}(t) = \sum_{mn}\sum_S P_S\, \beta_m^{(S)}\beta_n^{(S)*}\, |m\rangle\langle n|$$

Now, compare this last expression with Equation 2.16.4 to see that the matrix of the density operator (i.e., the density matrix) is

$$\rho_{mn} = \langle m|\hat{\rho}|n\rangle = \sum_S P_S\, \beta_m^{(S)}\beta_n^{(S)*} = \left\langle \beta_m^{(S)}\beta_n^{(S)*}\right\rangle_e$$

Whereas the density operator $\hat{\rho}$ gives the ensemble average of the wavefuntion projection operator

$$\overline{|\psi\rangle\langle\psi|} = \langle\ |\psi\rangle\langle\psi|\ \rangle_e$$

the density matrix element ρ_{mn} provides the ensemble average of the wavefunction coefficients

$$\overline{\beta_m^{(S)}\beta_n^{(S)*}} = \left\langle \beta_m^{(S)}\beta_n^{(S)*}\right\rangle_e$$

As a check on the previous calculation, let's find the density matrix directly from the definition of the density operator.

$$\hat{\rho}(t) = \sum_S P_S|\psi_S(t)\rangle\langle\psi_S(t)|$$

Operate with the unit projector $\langle a|$ and the unit ket $|b\rangle$ on both sides of the previous equation. The density matrix is

$$\langle a|\hat{\rho}(t)|b\rangle = \sum_S P_S \langle a|\psi_S(t)\rangle\langle\psi_S(t)|b\rangle \qquad (2.16.8)$$

A wavefuntion projected onto one of the basis vectors provides the corresponding Fourier expansion coefficient. Equation 2.17.5 provides

$$\langle a|\psi_S(t)\rangle = \sum_m \beta_m^{(S)}(t)\langle a|m\rangle = \beta_a^{(S)}(t)$$

by orthonormality of the wavefunctions (etc.). Therefore the density matrix in Equation 2.17.8 can be rewritten as

$$\langle a|\hat{\rho}(t)|b\rangle = \sum_S P_S \beta_a(t)\beta_b^*(t)$$

Topic 2.16.3: Ensemble and Quantum Mechanical Averages

The whole point of the density operator is to simultaneously provide two averages. The quantum mechanical average is used to find quantities such as average energy, electric field or position using only the quantum mechanical state of a given system. The ensemble average takes into account non-quantum mechanical influences such as variation in container size, slight differences in environment which are represented through a probability P_S. Notice in the definition of density operator

$$\hat{\rho}(t) = \sum_S P_S |\psi_S(t)\rangle\langle\psi_S(t)| \qquad (2.16.9)$$

that if one of the systems occurs at the exclusion of all others (say S=1)

$$\hat{\rho}(t) = |\psi_1(t)\rangle\langle\psi_1(t)| = |\psi(t)\rangle\langle\psi(t)| \qquad (2.16.10)$$

then the density operator only provides quantum mechanical averages; there is no macroscopic variation between systems to cause the wavefunction to vary from one system to the next. Density operators of the form 2.16.10 are called "pure" states of the system. Notice also that a "density operator" is also called a "state" which might be confused sometimes with calling the wave function $|\psi(t)\rangle$ a state. In the case of Equation 2.16.10, the density operator and the wave function are equivalent descriptions of the single quantum mechanical system and both obey a Schrodinger equation.

Define the quantum mechanical "q" and ensemble "e" averages for an operator $\hat{O}$ as follows

Quantum Mechanical	Ensemble
$\langle\hat{O}\rangle_q = \langle\psi|\hat{O}|\psi\rangle$	$\langle\hat{O}\rangle_e = \sum_S P_S \hat{O}_S$

where $|\psi\rangle$ is a typical quantum mechanical wave function. In what follows, the operator that appears in the ensemble average in the table above is just a number that depends on the particular system S (for example, it might be the system temperature that varies from one system to the next).

The discussion now demonstrates that the ensemble and quantum mechanical average of an operator $\hat{O}$ can be calculated as

$$\langle\langle\hat{O}\rangle\rangle = \mathrm{Tr}(\hat{\rho}\hat{O})$$

where for now, the double average symbol is meant as a reminder of the two types of averages. Recall the definition of trace,

$$\mathrm{Tr}(\hat{\rho}\hat{O}) = \sum_n \langle n|\hat{\rho}\hat{O}|n\rangle$$

where the summation is over the basis set $\{|n\rangle = |u_n\rangle\}$ which is usually the set of energy eigenstates. Recall that the trace is independent of the particular basis set chosen; that is, any basis set for the vector space can be chosen to evaluate the trace and the result is the same. Evaluating the last expression in terms of matrices by inserting the closure relation between the two operators provides

$$\mathrm{Tr}(\hat{\rho}\hat{O}) = \sum_n \langle n|\hat{\rho}\left[\sum_m |m\rangle\langle m|\right]\hat{O}|n\rangle = \sum_{m,n}\langle n|\hat{\rho}|m\rangle\langle m|\hat{O}|n\rangle = \sum_{m,n}\rho_{nm}O_{mn} \qquad (2.16.11)$$

The following work looks for an equation very similar to the previous one.

First, find the quantum mechanical average of an operator for the specific system S starting with

$$\langle\hat{O}\rangle_q^{(S)} = \langle\psi_S|\hat{O}|\psi_S\rangle$$

where

$$|\psi_S(t)\rangle = \sum_n \beta_n^{(S)}(t)|u_n\rangle$$

is the wave function for the system S. Substituting the wave function into the operator expression provides

$$\langle\hat{O}\rangle_q^{(S)} = \left\{\sum_n \beta_n^{*(S)}\langle u_n|\right\}\hat{O}\left\{\sum_m \beta_m^{(S)}(t)|u_m\rangle\right\} = \sum_{nm}\beta_n^{*(S)}\beta_m^{(S)}\langle u_n|\hat{O}|u_m\rangle = \sum_{mn}\beta_n^{*(S)}\beta_m^{(S)}O_{nm}$$

There is one such average for each system S since there is a different wave function for each system. For a given system S, this last expression gives the quantum mechanical average of the operator for that one system.

Now find the ensemble average of the quantum mechanical expectation value of the operator $\hat{O}$. Define the double-average notation and write

$$\langle\langle\hat{O}\rangle_q\rangle_e = \langle\langle\hat{O}\rangle\rangle = \left\langle\sum_{nm}\beta_m^{(S)}\beta_n^{*(S)}O_{nm}\right\rangle_e = \sum_{nm}\langle\beta_m^{(S)}\beta_n^{*(S)}\rangle_e O_{nm} = \sum_{nm}\left(\sum_S P_S\,\beta_m^{(S)}\beta_n^{*(S)}\right)O_{nm} = \sum_{nm}\rho_{nm}O_{nm}$$

Comparing this last result with Equation 2.16.11, we find

$$\langle\langle\hat{O}\rangle\rangle = \mathrm{Tr}(\hat{\rho}\hat{O}) \qquad (2.16.12)$$

Most books leave off the double brackets and just write

$$\langle\hat{O}\rangle = \mathrm{Tr}(\hat{\rho}\hat{O}) \qquad (2.16.13)$$

One can see for a pure state with $\hat{\rho} = |\psi(t)\rangle\langle\psi(t)|$ that

$$\langle \hat{O} \rangle = \mathrm{Tr}(\hat{\rho}\hat{O}) = \mathrm{Tr}\{|\psi(t)\rangle\langle\psi(t)|\hat{O}\} = \sum_n \langle u_n|\psi(t)\rangle\langle\psi(t)|\hat{O}|u_n\rangle$$

$$= \sum_n \langle\psi(t)|\hat{O}|u_n\rangle\langle u_n|\psi(t)\rangle = \langle\psi(t)|\hat{O}|\psi(t)\rangle$$

where the first summation uses the definition of trace and the last step uses the closure relation for the states $|n\rangle = |u_n\rangle$. For the pure state, the trace formula reduces to the ordinary quantum mechanical average of $\langle \hat{O} \rangle = \langle\psi(t)|\hat{O}|\psi(t)\rangle$.

Topic 2.16.4: Some Properties

(1) If $\hat{\rho} = |\psi\rangle\langle\psi|$ that is, there is no statistical mixture $P_\psi = 1$ so that $\hat{\rho}$ is in a pure state, then

$$\hat{\rho}\hat{\rho} = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = \hat{\rho}$$

so that $\hat{\rho}$ is idempotent. The only possible eigenvalues for this particular density operator are 0 and 1.

$$\hat{\rho}|v\rangle = v|v\rangle \rightarrow \hat{\rho}\hat{\rho}|v\rangle = v|v\rangle \rightarrow v^2|v\rangle = v|v\rangle \rightarrow v^2 = v \rightarrow v = 0,1$$

(2) All density operators are Hermitian

$$\hat{\rho}^+ = \left\{\sum_\psi P_\psi|\psi\rangle\langle\psi|\right\}^+ = \sum_\psi P_\psi\{|\psi\rangle\langle\psi|\}^+ = \sum_\psi P_\psi|\psi\rangle\langle\psi| = \hat{\rho}$$

since the probability is a real number.

(3) Diagonal elements are the probability that a system will be found in a specific eigenstate. The diagonal elements take into account both ensemble and quantum mechanical probabilities. Let $\{|a\rangle\}$ be a complete set of states (basis states) and let the wavefunction for each system have the form

$$|\psi(t)\rangle = \sum_a \beta_a^{(\psi)}(t)|a\rangle$$

The diagnonal elements of the density matrix are then

$$\rho_{aa} = \langle a|\hat{\rho}|a\rangle = \langle a|\left\{\sum_\psi P_\psi|\psi\rangle\langle\psi|\right\}|a\rangle = \sum_\psi P_\psi\langle a|\psi\rangle\langle\psi|a\rangle = \sum_\psi P_\psi\beta_a^{*(\psi)}\beta_a^{(\psi)} = \overline{|\beta_a|^2} = \overline{\mathrm{prob(a)}}$$

(4) The sum of the diagonal elements must be unity.

$$\mathrm{Tr}(\hat{\rho}) = \sum_n \rho_{nn} = 1$$

since the diagonal elements are probabilities and all possible states correspond to diagnol elements.

Topic 2.16.5: Intuitive Examples

(1) What is the initial density operator $\hat{\rho}(0)$ for the two-level atoms given in the table?

Initial Wave Function $\lvert\psi_S(0)\rangle$	Probability P_s
$\lvert\psi_1\rangle = 0.98\lvert u_1\rangle + 0.19\lvert u_2\rangle$	1/4
$\lvert\psi_1\rangle = 0.90\lvert u_1\rangle + 0.43\lvert u_2\rangle$	1/2
$\lvert\psi_1\rangle = 0.95\lvert u_1\rangle + 0.31\lvert u_2\rangle$	1/4

The initial density operator is given by

$$\hat{\rho}(0) = \sum_{S=1}^{3} P_S \lvert\psi_S(0)\rangle\langle\psi_S(0)\rvert$$

Substituting the probabilities and intial wave functions provides

$$\begin{aligned}\hat{\rho}_S(0) &= P_1\lvert\psi_1(0)\rangle\langle\psi_1(0)\rvert + P_2\lvert\psi_2(0)\rangle\langle\psi_2(0)\rvert + P_3\lvert\psi_3(0)\rangle\langle\psi_3(0)\rvert \\ &= \frac{1}{4}\left[0.98\lvert u_1\rangle + 0.19\lvert u_2\rangle\right]\left[0.98\langle u_1\rvert + 0.19\langle u_2\rvert\right] + \\ &+ \frac{1}{2}\left[0.90\lvert u_1\rangle + 0.43\lvert u_2\rangle\right]\left[0.90\langle u_1\rvert + 0.43\langle u_2\rvert\right] + \\ &+ \frac{1}{4}\left[0.95\lvert u_1\rangle + 0.31\lvert u_2\rangle\right]\left[0.95\langle u_1\rvert + 0.31\langle u_2\rvert\right]\end{aligned}$$

Collecting terms

$$\begin{aligned}\hat{\rho} = &(0.24 + 0.4 + 0.22)\lvert u_1\rangle\langle u_1\rvert + (0.046 + 0.19 + 0.071)\lvert u_1\rangle\langle u_2\rvert + \\ &+ (0.046 + 0.19 + 0.071)\lvert u_2\rangle\langle u_1\rvert + (0.009 + 0.092 + 0.022)\lvert u_2\rangle\langle u_2\rvert\end{aligned}$$

or finally,

$$\hat{\rho}(0) = 0.86\lvert u_1\rangle\langle u_1\rvert + 0.307\lvert u_1\rangle\langle u_2\rvert + 0.307\lvert u_2\rangle\langle u_1\rvert + 0.14\lvert u_2\rangle\langle u_2\rvert$$

Notice how the coefficients of the first and last term add to one... this is no accident. For future reference, the *density matrix* is

$$\underline{\rho} = \begin{bmatrix} 0.86 & 0.307 \\ 0.307 & 0.14 \end{bmatrix}$$

in the basis set $\{\lvert u_1\rangle, \lvert u_2\rangle\}$. As discussed later, the diagonal elements of the density matrix correspond to the probability that a particle will be found in the level $\lvert u_1\rangle, \lvert u_2\rangle$.

(2) Let $\hat{H}$ be the Hamiltonian for a two-level system with energy eigenvectors $\{\lvert u_1\rangle, \lvert u_2\rangle\}$ so that $\hat{H}\lvert u_1\rangle = E_1\lvert u_1\rangle$ and $\hat{H}\lvert u_2\rangle = E_2\lvert u_2\rangle$. What is the matrix of $\hat{H}$ with respect to the basis vectors $\{\lvert u_1\rangle, \lvert u_2\rangle\}$?

The matrix elements of $\hat{H}$ are found as

$$H_{ab} = \langle u_a | \hat{H} | u_b \rangle = E_b \delta_{ab}$$

which can be written as

$$\underline{H} = \begin{bmatrix} E_1 & 0 \\ 0 & E_2 \end{bmatrix}$$

(3) What is the ensemble-averaged energy $\langle \hat{H} \rangle \equiv \langle\langle \hat{H} \rangle\rangle$? Assume all of the information is the same as for examples one and two.

It is appropriate to evaluate the average given by

$$\langle \hat{H} \rangle = \mathrm{Tr}(\hat{\rho}\hat{H})$$

Insert the basis vectors as required by the trace and then insert the closure relation between the two operators. The resulting formula is identical to taking the trace of the product of two matrices.

$$\mathrm{Tr}(\hat{\rho}\hat{H}) = \mathrm{Tr}(\underline{\rho}\underline{H}) = \mathrm{Tr}\begin{bmatrix} 0.86 & 0.307 \\ 0.307 & 0.14 \end{bmatrix}\begin{bmatrix} E_1 & 0 \\ 0 & E_2 \end{bmatrix} = \mathrm{Tr}\begin{bmatrix} 0.86E_1 & 0.307E_1 \\ 0.307E_2 & 0.14E_2 \end{bmatrix}$$

Of course, in switching from operators to matrices, the isomorphism between operators and matrices is invoked. Operations using the operators must be equivalent to operations using the corresponding matrices. The trace of a matrix is just the sum of the diagonal elements

$$\langle \hat{H} \rangle = \mathrm{Tr}\,(\hat{\rho}\hat{H}) = 0.86E_1 + 0.14E_2$$

So the average is no longer equal to the eigenvalue E_1 or E_2. The average energy represents a combination of the energies dictated by both the quantum mechanical and ensemble probabilities.

(4) What is the probability that an electron will be found in the state $|u_1\rangle$? Assume that the information from Examples 1-3 is valid. We have not discussed this use of the density matrix yet.

We are assuming that the density matrix is given by

$$\underline{\rho} = \begin{bmatrix} 0.86 & 0.307 \\ 0.307 & 0.14 \end{bmatrix}$$

The answer is $\mathrm{prob}(\#1) = \langle u_1 | \hat{\rho} | u_1 \rangle = \rho_{11} = 0.86$. In fact, the probability of the first state being occupied can be found directly from the definition of the density operator

$$\langle 1|\hat{\rho}|1\rangle = \langle 1 \left| \left[\sum_s P_s |\psi_s\rangle\langle\psi_s| \right] \right| 1\rangle = \sum_s P_s \langle 1|\psi_s\rangle\langle\psi_s|1\rangle = \sum_s P_s \beta_1^{(s)} \beta_1^{(s)*} = \overline{\beta_1 \beta_1^*}$$

(5) Example of Diagonal Elements for Two-Level Atoms

Suppose a system contains N independent two-level atoms (per unit volume). A physical interpretation can be given to the diagonal and off-diagonal elements of the density matrix.

Let ρ_{mn} is the density matrix for an ensemble of two-level atoms. The number of atoms (per unit volume) in state $|a\rangle$ is given by

$$N_a = (\text{total number})\,(\text{Prob of state a}) = N\,\rho_{aa} \quad (2.16.14)$$

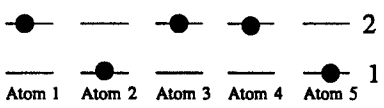


Figure 2.16.7: Five independent two-level atoms. Each one has the characteristics of the ensemble. $\rho_{11} = 2/5$ and $\rho_{22} = 3/5$

The diagonal elements of the density matrix can be calculated by assuming that the wavefunctions $|\psi_S\rangle$ can only be either $|u_1\rangle$ or $|u_2\rangle$. The density operator has the form

$$\hat{\rho} = \sum_{S=1}^{2} P_S |\psi_S\rangle\langle\psi_S| = P_1|u_1\rangle\langle u_1| + P_2|u_2\rangle\langle u_2| + \cdots$$

or, equivalently, the matrix is $\rho_{aa} = \langle u_a|\hat{\rho}|u_a\rangle \rightarrow \underline{\rho} = \begin{bmatrix} P_1 & \cdots \\ \cdots & P_2 \end{bmatrix}$. The figure clearly shows that $\text{Prob}(1) = P_1 = 3/5$ and $\text{Prob}(2) = P_2 = 2/5$. An infinite number of possible systems has been reduced to just two since owing to the assignment of probabilities.

(6) What if we had defined the occupation number operator $\hat{n}$ to be

$$\hat{n}|1\rangle = 1|1\rangle \qquad \hat{n}|2\rangle = 2|2\rangle$$

Calculate the expectation value of $\hat{n}$ using the trace formula for the density operator. For this exercise, assume the off-diagonal elements of the density matrix are zero.

$$\langle\hat{n}\rangle = \text{Tr}(\hat{\rho}\hat{n}) = \text{Tr}\begin{bmatrix} 2/5 & 0 \\ 0 & 3/5 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = \text{Tr}\begin{bmatrix} 2/5 & 0 \\ 0 & 6/5 \end{bmatrix} = \frac{2}{5} + \frac{6}{5} = \frac{8}{5}$$

This just says that the average state is somewhere between "1" and "2". Looking at the figure, the average state should be

$$1\cdot\text{Prob}(1) + 2\cdot\text{Prob}(2) = 1\frac{2}{5} + 2\frac{3}{5} = \frac{8}{5}$$

as found with the density matrix. Notice that making the off-diagonal elements of the density matrix has no effect on this calculation.

Topic 2.16.6: Off-Diagonal Elements

The off-diagonal elements of the density matrix are similar to the probability amplitude that a particle is simultaneously in states "n" *and* "m" for example.

$$\rho_{mn} = \langle m|\hat{\rho}|n\rangle \sim \langle m|\psi\rangle\langle\psi|n\rangle \sim \psi_m(x)\,\psi_n^*(x)$$

where ψ_m is the m^{th} eigenfunction of the Hamiltonian (for example). The probability (density) is then the modulus squared of the matrix element.

As a more physical illustration, the off-diagonal elements are related to the polarization and the expected dipole matrix elements. The dipole moment is related to the allowed transitions for an atomic system.

The dipole operator is defined by

$$\vec{\mu} = e\vec{r}$$

where e>0 is the basic electronic charge. Assume the coordinate representation. The are permanent and induced dipole moments. The permanent dipole moment refers to the dipole moment of an atom without an applied perturbing electromagnetic field.

$$\mu_{nn} = \langle n|\mu|n\rangle$$

where the states $|n\rangle$ are the energy eigenstates of the atom. Usually, for simplicity, it is assumed that the permanent dipole moment is zero. This is easy to see for spherically symmetrical wavefunctions by calculating the expected dipole moment as

$$\langle a|\hat{\mu}_x|a\rangle = e\langle u_a|x|u_a\rangle = e\int_{-\infty}^{\infty} dx \;\; \underset{\leftarrow \text{odd} \rightarrow}{u_a^*(x)\, x\, u_a(x)} = 0$$

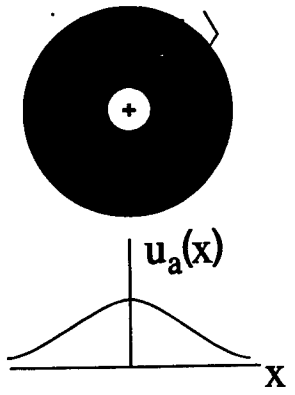


Figure 2.16.8: Spherically symmetric wavefunction

At any rate, even if the permanent dipole moment is not zero, it does not affect the transition rates that depend on the off-diagonal elements.

The "induced dipole moment" occurs when a perturbing electromagnetic field is present. The induced dipole moments are given by

$$\langle a|\hat{\mu}|b\rangle = \mu_{ab} \text{ with } a \neq b$$

and are related to the transition selection rules.

The following discussion leads to the relation between the off-diagonal density matrix elements and the polarization. First consider the definition of polarization. Let N be the number of atoms per unit volume.

$$\text{Polarization} = P = \frac{\text{Total Dipole Moment}}{\text{Vol}} = N\langle\mu\rangle = N\,\text{Tr}(\hat{\rho}\hat{\mu})$$

Assume a perturbing EM field that provides the interaction energy V so that

$$\hat{H} = \hat{H}_o + V \quad \text{and} \quad \hat{H}|n\rangle = E_n|n\rangle$$

and assume a two-level atom so that n=1,2. Only the energy eigenstates of the *unperturbed* Hamiltonian are required for this calculation. Using the isomorphism between operators and matrices provides

$$P = N\,\text{Tr}(\hat{\rho}\hat{\mu}) = N\,\text{Tr}\,\underline{\rho}\underline{\mu} = N\sum_a \left(\underline{\rho}\underline{\mu}\right)_{aa}$$

However, the parenthesis contains a product of matrices so

$$P = N\sum_{a}\left(\underline{\rho}\underline{\mu}\right)_{aa} = N\sum_{a,b=1}^{2}\rho_{ab}\mu_{ba} = N\{\rho_{12}\mu_{12} + \rho_{21}\mu_{21}\}$$

Assuming the lack of a permanent dipole (i.e., $\mu_{11} = \mu_{22} = 0$). If we assume (as is usually the case) that

$$\mu = \mu_{12} = \mu_{21} = \text{real}$$

then the equation for the polarization becomes

$$P = N\{\rho_{12}\mu_{12} + \rho_{21}\mu_{21}\} = N\mu\left(\rho_{12} + \rho_{21}\right)$$

So the off-diagonal elements of the density matrix are related to the polarization. We can go a step further by using the fact that the density operator is Hermitian.

$$\hat{\rho} = \hat{\rho}^{+} \rightarrow (\hat{\rho})_{12} = \left(\hat{\rho}^{+}\right)_{12} \rightarrow \rho_{12} = \rho_{21}^{*}$$

Now the polarization formula becomes

$$P = N\mu\left(\rho_{12} + \rho_{21}\right) = N\mu\left(\rho_{12} + \rho_{12}^{*}\right) = 2N\mu\,\mathrm{Re}\,\rho_{12}$$

In classical electromagnetics, it is the induced polarization that interacts with the EM field to generate or absorb EM energy. As discussed in a number of books (Yariv for example), this is the point of connection between classical EM and quantum theory.

Chapter 3: Quantized Fields

The study of light and EM fields through the centuries includes Newton's corpuscular theory, Young's interference experiments, Maxwell's Equations, Einstein and Planck's quanta, 20th century's field quantization, and Glauber and Yuen's coherent and squeezed optical states. This chapter explores the nature of the light and electromagnetic waves issuing from lasers, masers and RF transmitters in terms of quantum optics. Besides non-coherent EM waves, there are at least three especially important states including Fock, coherent and squeezed states. These states determine the noise levels. For example, shot noise is related to the coherent state while sub-shot noise signals (sub-poisson statistics) are related to the squeezed state. These states have equal application to massive particles (such as electrons) as to the massless photons. The three states of light are associated with the quantized Hamiltonian and electromagnetic fields which obtain through the quantization rules for their classical counterparts. This chapter begins the quest with a discussion of the classical vector potential and Gauge transformations. Next, the solutions to the optical schrodinger equation are discussed with special emphasis on the Fock, coherent and squeezed states. The quantization of the electromagnetic field is similar to the second quantization of the massive particle field. The formalism is applied to find the rate of stimulated and spontaneous emission. The chapter introduces the Wigner function.

* The generic term "light" is used for the photon field regardless of frequency.*

Section 3.1: A Brief Overview of Quantum Light

As is well known, atoms can emit light that is coherent with a driving optical field (stimulated emission) and they can also emit light on their own without a driving field (spontaneous emission). The spontaneous emission arises as a result of quantum vacuum fluctuations. To discuss vacuum fluctuations, an introduction to quantum electrodynamics is in order.

One of the most fundamental notions of Quantum ElectroDynamics (QED) is that of a *Fock* state. For a given spatial volume, Maxwell's ElectroMagnetic (EM) equations can be solved so as to define a set of allowed EM modes (for a cubic volume, the modes are sines and cosines). For Fock states, QED keeps track of the number of photons in a mode, which is characterized by the allowed wavelengths and polarization. In some sense, the modes appear as a framework or stage that the photons occupy. A Fock state is a mode with a definite number of photons (this means that each mode has a definite average power). In *classical* electrodynamics, a state without any photons corresponds to a mode without any amplitude. In QED, a state without any photons (the vacuum state) has an *average* electric field of zero, but the variance (which is proportional to the square of the field) is non-zero. This means that the value of the electric field can fluctuate away from the average of zero. As is well known, the non-zero variance refers to quantum vacuum fluctuations or noise; the vacuum state has the minimum quantum noise. There is a slight complication for engineering purposes however: the average electric field is zero for all Fock states and in general, the noise is greater than the minimum value set by

the vacuum. In addition, although the exact number of photons is known, the phase of the collection is not known.

A *coherent* state is a state with a nonzero average electric field and a fairly well defined phase. These are the states for which the electric fields can be pictured as sine and cosine waves. Coherent states can be produced by lasers, masers and some classical sources. The coherent state is actually a linear combination of all Fock states (the two types of states are seen to be quite different). One of the most important distinctions is that, for a coherent state with a given amplitude, the number of photons in the mode is described by a Poisson probability distribution (shot noise). The average number of particles is $\langle n \rangle$ while the standard deviation $\sigma=\sqrt{\langle n \rangle}$. For example, a beam with an average of $\langle n \rangle=100$ photons, will have a standard deviation of $\sqrt{\langle n \rangle}=10$ photons. One might reasonably expect the number of photons to range from 80 to 120 (almost 50% variation). This means that the number of photons in a beam can fluctuate such that the variance of the number is approximately equal to the average number of photons in the beam (shot noise). Now returning to the amplitude and phase, it just so happens that the noise in the amplitude and phase of the coherent state is no worse than that of the vacuum state (regardless of the amplitude of the coherent state). A time dependent electric field can be decomposed into a time dependent sine and cosine term so that the uncertainty in the amplitudes of these quadrature terms can be specified as an alternative to specifying the uncertainty in the amplitude and phase.

A *squeezed* vacuum state is obtained from the quantum vacuum state by reducing the noise in one set of parameters while adding it to another (i.e., "squeezing the noise out"). Squeezing the vacuum is equivalent to squeezing the coherent state since the vacuum and coherent states have the same type and amount of noise. For example, noise can be removed from one quadrature term of the electric field (quadrature squeezing) but that removed noise reappears in the other quadrature term. Similarly, a "quiet" photon stream (i.e. a number squeezed state) obtains when noise is removed from the photon-number but reappears in the phase; a quiet photon stream is describe by "sub-Poisson" statistics. Similarly, there are phase squeezed states. Squeezed coherent states can be produced, detected and used for low noise applications. Figure 3.1.1 shows examples of laser light moving past an observer. The top portion shows "coherent light" (i.e., light in a coherent state) where the number of photons in equal beam-lengths can vary from one length-interval to the next. The number of photons follows the Poisson probability distribution. The bottom figure shows a "number squeezed state" where the equal lengths have equal numbers of photons. Apparently, a number-squeezed state is equivalent to a Fock state.

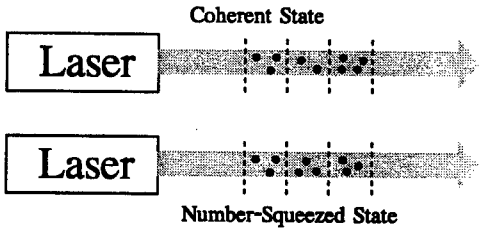


Figure 3.1.1: Pictorial representation of the coherent and number-squeezed states.

Spontaneous emission (RIN) is another form of noise in the laser although it is certainly not termed "noise" for a Light Emitting Diode (LED). As is well known, spontaneous emission in a laser is necessary to start the laser oscillation but, in addition to larger than necessary threshold current, it also wastes energy. Further, it partly determines the ability of the laser to produce squeezed light as well as the power levels to produce large S/N and dynamic range. An interesting fact about spontaneous emission is

that it is a result of quantum vacuum fluctuations and that it is not a property of the emitting collection of atoms. The spontaneous emission is initiated by the fluctuating electric field of the vacuum state. The rate of spontaneous emission can be modified by changing the number of accessible vacuum modes (there is one vacuum mode for each wavelength and polarization allowed by the boundary conditions on the enclosed volume) coupled to the atomic ensemble. The relatively new field of Cavity QED describes the theory and measurement of both spontaneous and stimulated emission for which these unusual cavity effects become important. To characterize the effect of spontaneous emission on another laser or device, it is necessary to understand the effects of vacuum entropy.

There is a nice conceptual aspect of the theory of quantized fields. The theory of quantized fields mathematically unifies the picture of the EM field as both a particle and a wave. The quantized electric field consists of a wave portion (described by the complex traveling wave) and a particle portion described by creation and annihilation operators. Both the wave and operators appear in the expression for the quantized field. Quantum field theory mathematically unifies the wave and particle pictures for matter not just photons or electrons.

The previous few paragraphs point out the importance of quantum optics and some very interesting sources of noise in electromagnetic systems. All forms of noise such as RIN and thermal noise need to be addressed.

Section 3.2: Coulomb Gauge, Classical Vector Potential and the EM Fields

In classical electrodynamics, the magnetic and electric fields are physical quantities. The Hamiltonian for a system of photons is written in terms of the vector (and scalar) potential. Substituting operators for the classical dynamical variables quantizes the Hamiltonian for the electromagnetic fields. The electric and magnetic fields are commonly derived from a vector potential. The relation between the fields and the potentials assumes a particular gauge transformation. A gauge transformation is a change made to the potentials that leaves the fields unchanged. As a result, Maxwell's equations are also invariant with respect to Gauge transformations. Sometimes the Coulomb gauge is also called the transverse gauge because the direction of the vector potential is perpendicular to the propagation wavevector. The Lorentz gauge manifests the relativistic invariance of the electromagnetic theory. Although sometimes the fields are thought of as the "real" objects of interest and the vector potential is just a mathematical construction, the potentials have real affects for device engineering. The Aharanov-Bohm devices provide an example where the vector potential (and not the fields) is used for modulating currents. The vector potential is particularly simple to quantize in the Coulomb gauge but the resulting equations are not relativistically invariant. This section focuses on the Coulomb gauge.

Topic 3.2.1: Magnetic and Electric Fields in the Coulomb Gauge

This topic shows (1) the relations between the fields and potentials for the Coulomb gauge and (2) the fields satisfy Maxwell's equations. The next topic specializes to a source-free region of space and a subsequent topic indicates how the Coulomb gauge arises.

In the Coulomb guage, the magnetic and electric fields are related to the vector potential $\vec{A}(\vec{r},t)$ and the scalar potential "V" by

$$\vec{B} = \nabla x \vec{A}(\vec{r},t) \qquad \text{and} \qquad \vec{E} = -\frac{\partial \vec{A}(\vec{r},t)}{\partial t} - \nabla V \tag{3.1.1}$$

The Coulomb gauge consists of those vector potentials $\vec{A}(\vec{r},t)$ that satisfy the Coulomb condition $\nabla \cdot \vec{A} = 0$. This section shows that the scalar potential is the "electrostatic voltage due to charges" and that the vector potential can pictured as a traveling wave when the electric and magnetic fields are traveling waves.

The magnetic and electric fields derived from the potentials in the Coulomb gauge satisfy Maxwell's equations.

$$\nabla \cdot \vec{B} = 0 \qquad \nabla x \vec{E} = -\frac{\partial \vec{B}}{\partial t}$$

$$\nabla \cdot \vec{D} = \rho \qquad \nabla x \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t} \tag{3.1.2}$$

where $\vec{D} = \varepsilon \vec{E} = \varepsilon_o \vec{E} + \vec{P} = \varepsilon_o \vec{E} + \varepsilon_o \chi \vec{E} = \varepsilon_o (1+\chi)\vec{E}$, $\vec{B} = \mu \vec{H} = \mu_o \vec{H}$ and where $\varepsilon_o, P, \chi, \mu_o$ are the permittivity of free-space, polarization, susceptibility, and

permeability of free-space. We assume an isotropic, homogeneous, non-magnetic medium.

(1) First show that the magnetic field derived from the vector potential satisfies $\nabla \cdot \vec{B} = 0$. The curl of the vector potential is defined by the determinant

$$\vec{B} = \nabla x \vec{A} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \partial_x & \partial_y & \partial_z \\ A_x & A_y & A_z \end{vmatrix}$$

and the divergence of the magnetic field gives the triple product

$$\nabla \cdot \vec{B} = \nabla \cdot \nabla x \vec{A} = \begin{vmatrix} \partial_x & \partial_y & \partial_z \\ \partial_x & \partial_y & \partial_z \\ A_x & A_y & A_z \end{vmatrix} = \partial_x \left(\partial_y A_z - \partial_z A_y \right) - \partial_y \left(\partial_x A_z - \partial_z A_x \right) + \partial_z \left(\partial_x A_y - \partial_y A_x \right) = 0$$

(2) Next demonstrate that E,B derived from the vector potential satisfy

$$\nabla x \vec{E} = -\frac{\partial \vec{B}}{\partial t}$$

Starting with $\nabla \times \vec{E}$ and substituting for the electric field using $\vec{E} = -\frac{\partial \vec{A}(\vec{r},t)}{\partial t} - \nabla V$ gives

$$\nabla \times \vec{E} = \nabla \times \left(-\frac{\partial \vec{A}}{\partial t} - \nabla V \right) = -\frac{\partial}{\partial t} \nabla \times \vec{A} - \nabla \times \nabla V = -\frac{\partial}{\partial t} \vec{B}$$

which is the desired Maxwell equation. The final step requires the definition $\vec{B} = \nabla x \vec{A}(\vec{r},t)$ (in the Coulomb gauge) and the fact that curl of a gradient is always zero $\nabla \times \nabla V = 0$.

(3) The third of Maxwell's equations

$$\nabla \cdot \vec{D} = \rho$$

is satisfied by the fields derived from the potentials. The scalar potential "V" is the important quantitiy and yields Poisson's equation. Substituting for the electric field gives

$$\frac{\rho}{\varepsilon_o} = \nabla \cdot \vec{E} = \nabla \cdot \left(-\frac{\partial \vec{A}}{\partial t} - \nabla V \right) = -\frac{\partial}{\partial t} \nabla \cdot \vec{A} - \nabla^2 V = -\nabla^2 V$$

where the last step follows from the Coulomb gauge condition. The last equation is Poisson's equation for the voltage V

$$\nabla^2 V = -\frac{\rho}{\varepsilon_o}$$

(4) The final Maxwell equation provides a wave equation for the vector potential. Starting with

$$\nabla \times \vec{B} = \varepsilon \mu \frac{\partial \vec{E}}{\partial t} + \vec{J}$$

Substituting for E,B from potentials

$$\vec{B} = \nabla \times \vec{A}(\vec{r},t) \qquad \text{and} \qquad \vec{E} = -\frac{\partial \vec{A}}{\partial t} - \nabla V$$

gives

$$\begin{aligned} \nabla \times \nabla \times \vec{A} &= \varepsilon\mu \frac{\partial}{\partial t}\left(-\frac{\partial \vec{A}}{\partial t} - \nabla V \right) + \vec{J} \\ &= -\varepsilon\mu \frac{\partial^2 \vec{A}}{\partial t^2} - \varepsilon\mu \frac{\partial}{\partial t} \nabla V + \vec{J} \end{aligned} \tag{3.1.3}$$

The double cross product is evaluated using the differential form of the "BAC-CAB" rule which gives

$$\nabla \times \nabla \times \vec{A} = \nabla\left(\nabla \cdot \vec{A}\right) - \nabla^2 \vec{A}$$

The first term on the right hand side is zero because of the Coulomb gauge condition $\nabla \cdot \vec{A} = 0$. Therefore, Equation 3.1.3 becomes a wave equation

$$\nabla^2 \vec{A} - \varepsilon\mu \frac{\partial^2 \vec{A}}{\partial t^2} = +\varepsilon\mu \frac{\partial}{\partial t} \nabla V - \vec{J} \tag{3.1.4}$$

where the speed of light in the medium is

$$v = \sqrt{\frac{1}{\mu\varepsilon}}$$

Interestingly, the scalar potential "V" is related to the "instantaneously" propagating longitudinal component of the vector potential (i.e., component of A parallel to k).

Topic 3.2.2: The Fields in a Source-Free Region of Space

In a source-free region of space, the charge and current density are zero

$$\rho = 0 \qquad \vec{J} = 0$$

The magnetic and electric fields in the Coulomb gauge become

$$\vec{B} = \nabla x \vec{A}(\vec{r},t) \qquad \text{and} \qquad \vec{E} = -\frac{\partial \vec{A}(\vec{r},t)}{\partial t}$$

assuming that the source-free voltage is zero so that $\nabla V = 0$. The wave equation for the vector potential given by Equation 3.1.4, with the source term set to zero

$$\varepsilon\mu \frac{\partial}{\partial t} \nabla V - \vec{J} = 0$$

becomes

$$\nabla^2 \vec{A} - \varepsilon\mu \frac{\partial^2 \vec{A}}{\partial t^2} = 0 \tag{3.1.5}$$

where, again, the speed of light in the medium is

$$v = \sqrt{\frac{1}{\mu\varepsilon}}$$

Example: Most readers are familiar with the vector potential, but this is an example just to provide a picture. Find $\vec{E}$, $\vec{B}$, and an alternate expression for $\nabla\cdot\vec{A}=0$ using the vector potential given by

$$\vec{A}=\frac{\vec{A}_o}{\omega}e^{i(\vec{k}\cdot\vec{r}-\omega t)}$$

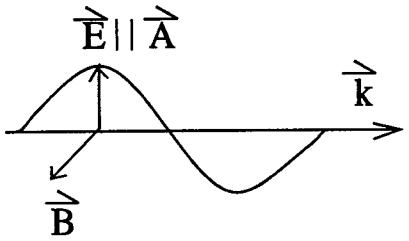


Figure 3.2.1: Vector potential in the Coulomb Gauge.

where $\vec{k}=k\hat{z}$.

Solution: The electric field is

$$\vec{E}=-\frac{\partial\vec{A}}{\partial t}=i\vec{A}_o e^{i(kz-\omega t)} \qquad (3.1.6)$$

The magnetic field is

$$\vec{B}=\nabla\times\vec{A}=\frac{i\vec{k}\times\vec{A}_o}{\omega}e^{i(\vec{k}\cdot\vec{r}-\omega t)}=i\frac{k}{\omega}\hat{z}\times\vec{A}_o e^{i(\vec{k}\cdot\vec{r}-\omega t)}=i\frac{\hat{z}\times\vec{A}_o}{c}e^{i(\vec{k}\cdot\vec{r}-\omega t)}$$

Finally, an alternate form for the gauge condition follows by substituting the vector potential into $\nabla\cdot\vec{A}=0$

$$0=\nabla\cdot\vec{A}=\frac{\vec{A}_o}{\omega}\nabla\cdot e^{i(\vec{k}\cdot\vec{r}-\omega t)}=\frac{\vec{k}\cdot\vec{A}}{\omega}$$

to give

$$\vec{k}\cdot\vec{A}=0 \qquad (3.1.7)$$

which is the reason the Coulomb gauge is sometimes called the transverse gauge. The vector part of $\vec{A}$ is perpendicular to the direction of propagation of the wave according to Equation 3.1.7 and it is parallel to the electric field according to Equation 3.1.6.

Topic 3.2.3: The Origin of the Coulomb Gauge Transformation

The gauge transformations all follow from an arbitrariness in the definition of the potentials (for example, everyone is familiar with the fact that the zero of the electric potential V can be shifted without affecting the fields). The Coulomb condition originates by requiring the potential V to satisfy Gauss' Law

$$\nabla\cdot\vec{E}=\frac{\rho}{\varepsilon_o}$$

by substituting $\vec{E}=-\frac{\partial\vec{A}(\vec{r},t)}{\partial t}-\nabla V$ from Equation 3.1.1.

$$\nabla\cdot\left[-\frac{\partial\vec{A}(\vec{r},t)}{\partial t}-\nabla V\right]=\frac{\rho}{\varepsilon_o}$$

The first term contains $\nabla\cdot\vec{A}$ which must be zero for V to be the voltage that satisfies Poissons' equation

$$\frac{\rho}{\varepsilon_o}=-\nabla\cdot\nabla V=-\nabla^2 V$$

Section 3.3: The Basis Vector Expansion of the Vector Potential

The vector potential is of primary interest because the other fields can be derived from it and the energy can be written in terms of it. For the quantum optical field theory, Maxwell's equations with appropriate boundary conditions are solved for the physical system of interest using separation of variables and generalized Fourier expansions. The field is quantized by replacing the Fourier amplitudes with operators. The plane wave decomposition of the vector potential is most commonly employed for the quantization procedure (Topic 3.3.1). Other basis vector decompositions are useful for the wide variety of optical devices (Topic 3.3.2). Subsequent sections sections demonstrate the quantization of the Electromagnetic (EM) Hamiltonian and the vector potential.

Topic 3.3.1: Plane Wave Expansion

The electromagnetic (EM) field can be derived from a vector potential in the Coulomb gauge

$$\nabla \cdot \vec{A} = 0 \tag{3.3.1}$$

according to

$$\vec{E} = -\frac{\partial \vec{A}(\vec{r},t)}{\partial t} \quad \text{and} \quad \vec{B} = \nabla \times \vec{A}(\vec{r},t) \tag{3.3.2}$$

in a source-free region of space where the scalar potential is zero. As a note, recall that the Coulomb gauge is also called the transverse gauge because the vector $\vec{A}$ is perpendicular to the wave-vector $\vec{k}$ and the vector $\vec{A}(\vec{r},t)$ can be pictured as a traveling wave. These fields derived from the vector potential are combined with the classical Hamiltonian in the next section. Either the vector potential is assumed to satisfy periodic boundary conditions or the wave is assumed to be confined to a finite region of space. Either way, the procedure uses the Fourier series as opposed to the Fourier integral. The characteristic volume can be made arbitrarily large at the end of the calculation if desired.

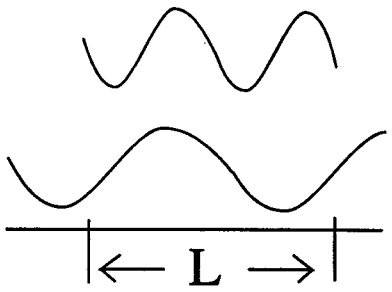


Figure 3.3.1: Two waves satisfying a periodic boundary condition over the length L.

The vector potential is a function that can be expanded in a basis set. The basis set for finite volumes V and discrete values of the wave vectors induced by the boundary conditions is

$$\left\{ \phi_{\vec{k}}(\vec{r}) = \frac{e^{i\vec{k}\cdot\vec{r}}}{\sqrt{V}} \right\} \tag{3.3.3}$$

where the wave vector $\vec{k}$ serves as the index. For a one-dimensional problem, the basis is $e^{ikx}/\sqrt{L}$ where the length "L" is also the longest allowed wavelength for periodic boundary conditions. If "L" is allowed to approach infinity, the 1-D basis set would be $e^{ikx}/\sqrt{2\pi}$ for the Fourier transform.

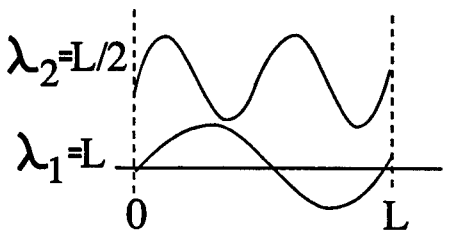


Figure 3.3.2: Waves confined to a finite length.

Periodic boundary conditions produce periodic basis functions that span a Hilbert space of periodic wave functions

$$\psi(x+L)=\psi(x)$$

as shown in Figure 3.3.2. Notice that the wavefunctions do not need to be zero at the boundaries. For the 3-D case, the volume V is related to L by $V=L^3$. The periodic boundary condition on the wave functions requires that the wave vectors $\vec{k}$ take on specific values. For the one-dimensional case, the allowed wavelengths for periodic boundary conditions are

$$\lambda_n=\frac{L}{n} \quad n=1,2,3...$$

and therefore the allowed wave vectors are

$$k_n=\frac{2\pi}{\lambda_n}=\frac{2\pi n}{L} \quad n=1,2,3... \tag{3.3.4}$$

For the three-dimensional case, each component of the wave vector $\vec{k}$ must satisfy an equation similar to Equations 3.3.4.

$$\vec{k}=\hat{x}\left(\frac{2\pi m}{L}\right)+\hat{y}\left(\frac{2\pi n}{L}\right)+\hat{z}\left(\frac{2\pi p}{L}\right) \quad m,n,p=1,2,3...$$

For light, the basic modes (i.e., the vectors in the basis set) are described by the allowed wavelengths (i.e., the allowed wave vectors or frequencies) and the polarization. For the most part, the polarization is ignored except possibly in final formulas.

Periodic vector potentials can be written as the Fourier sum

$$\vec{A}(\vec{r},t)=\frac{1}{\sqrt{\varepsilon_o}}\sum_{\vec{k}}\sqrt{\frac{\hbar}{2\omega_k}}\vec{A}_k(t)\frac{e^{i\vec{k}\cdot\vec{r}}}{\sqrt{V}} \tag{3.3.5}$$

for free space (the free space permittivity ε_o is replaced with ε for dielectric medium). The parameters

$$\sqrt{\frac{\hbar}{2\omega_k\varepsilon_o}}$$

are included for the MKS units. Assume that the angular frequency of the electromagnetic wave is always positive

$$\omega_k=\omega_{-k}>0$$

As an important point, the Fourier expansion coefficients must have the property that

$$\vec{A}_{-k}(t)=\vec{A}_k^*(t) \tag{3.3.6}$$

This is easy to demonstrate by noting that the vector potential must be *real* since it is related to observables E and B. Use Equation 3.3.5 and the relation

$$\vec{A}^*=\vec{A}$$

to obtain

$$\frac{1}{\sqrt{\varepsilon_o}}\sum_{\vec{k}}\sqrt{\frac{\hbar}{2\omega_k}}\vec{A}_k^*(t)\frac{e^{-i\vec{k}\cdot\vec{r}}}{\sqrt{V}}=\frac{1}{\sqrt{\varepsilon_o}}\sum_{\vec{k}}\sqrt{\frac{\hbar}{2\omega_k}}\vec{A}_k(t)\frac{e^{i\vec{k}\cdot\vec{r}}}{\sqrt{V}}$$

Replacing $\vec{k} \to -\vec{k}$ on the left-hand side of the previous relation and using $\omega_k = \omega_{-k}$ provides

$$\frac{1}{\sqrt{\varepsilon_o}}\sum_{\vec{k}}\sqrt{\frac{\hbar}{2\omega_k}}\vec{A}^*_{-k}(t)\frac{e^{+i\vec{k}\cdot\vec{r}}}{\sqrt{V}} = \frac{1}{\sqrt{\varepsilon_o}}\sum_{\vec{k}}\sqrt{\frac{\hbar}{2\omega_k}}\vec{A}_k(t)\frac{e^{i\vec{k}\cdot\vec{r}}}{\sqrt{V}}$$

Comparing both sides (i.e., using the orthonormality of the basis functions) gives the desired results

$$\vec{A}_{-k}(t) = \vec{A}^*_k(t)$$

The previous section shows that, in the Coulomb gauge, the vector potential must satisfy the wave equation

$$\nabla^2\vec{A} - \frac{1}{c^2}\frac{\partial^2\vec{A}}{\partial t^2} = 0 \qquad (3.3.7)$$

Substituting Equations 3.3.5 into Equation 3.3.7 requires two differential operators be calculated

$$\nabla^2\vec{A}(\vec{r},t) = \frac{1}{\sqrt{\varepsilon_o}}\sum_{\vec{k}}\sqrt{\frac{\hbar}{2\omega_k}}\vec{A}_k(t)\left(-k^2\right)\frac{e^{i\vec{k}\cdot\vec{r}}}{\sqrt{V}}$$

and

$$\frac{\partial^2}{\partial t^2}\vec{A}(\vec{r},t) = \frac{1}{\sqrt{\varepsilon_o}}\sum_{\vec{k}}\sqrt{\frac{\hbar}{2\omega_k}}\frac{\partial^2\vec{A}_k(t)}{\partial t^2}\frac{e^{i\vec{k}\cdot\vec{r}}}{\sqrt{V}}$$

So therefore, Equation 3.3.7 becomes

$$-k^2\vec{A}_{\vec{k}} - \frac{1}{c^2}\frac{\partial^2\vec{A}_{\vec{k}}}{\partial t^2} = 0$$

Using $c^2 = \omega^2/k^2$ gives

$$\frac{\partial^2\vec{A}_{\vec{k}}(t)}{\partial t^2} + \omega^2\vec{A}_{\vec{k}}(t) = 0 \qquad (3.3.8)$$

The Fourier transformed wave equation (Equation 3.3.8) has two solutions

$$A_k = b_k e^{-i\omega_k t} + a_k e^{+i\omega_k t} \qquad \omega_k > 0$$

where *the vector potential is considered first as a scalar quantity* A_k. Using the "reality" of the vector potential, or equivalently $A_{-k}(t) = A^*_k(t)$, gives

$$b^*_k e^{+i\omega_k t} + a^*_k e^{-i\omega_k t} = b_{-k}e^{-i\omega_k t} + a_{-k}e^{+i\omega_k t}$$

so that

$$b^*_k = a_{-k} \to a_k = b^*_{-k}$$

Therefore, the general solution to Equation 3.3.8 is

$$A_k = b_k e^{-i\omega_k t} + b^*_{-k}e^{+i\omega_k t} \qquad \omega_k > 0 \qquad (3.3.9)$$

Next include the vector part of $\vec{A}$. Returning to Equation 3.3.8, let $\tilde{e}_{ks}$ be unit vectors denoting the two possible directions of polarization (i.e., S=1,2) as shown in Figure 3.3.3. Recall the

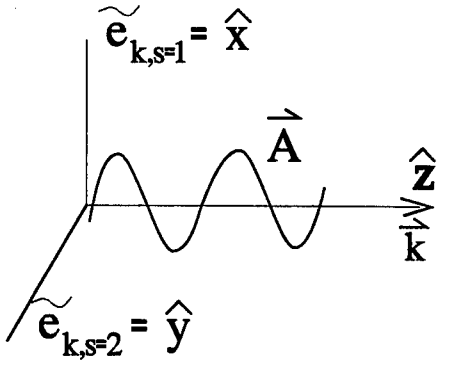


Figure 3.3.3: The two polarization modes.

polarization of the vector potential is parallel to the polarization of the electric field. There are two allowed polarizations of the electromagnetic fields that are consistent with the requirements of the Coulomb gauge (i.e., transverse gauge). The unit vectors $\tilde{e}_{ks}$ represent the $\hat{x}$ and $\hat{y}$ directions for a wave propagating along the $\hat{z}$ direction. The sum in Equation 3.3.5 is actually over all allowed modes and must include "s". For now, combine the k,s subscripts into the k subscript for simplicity. Writing Equation 3.3.9 in vector notation

$$\vec{A}_k = \vec{b}_k e^{-i\omega_k t} + \vec{b}^*_{-k} e^{+i\omega_k t} \qquad \omega_k > 0$$

or, setting $\vec{b}_k = \tilde{e}_k b_k$, the solution of Equations 3.3.8 is

$$\vec{A}_k = \tilde{e}_k b_k e^{-i\omega_k t} + \tilde{e}^*_{-k} b^*_{-k} e^{+i\omega_k t} \qquad \omega_k > 0$$

Now the vector potential can be written as

$$\vec{A}(\vec{r},t) = \frac{1}{\sqrt{\varepsilon_o}} \sum_{\vec{k}} \sqrt{\frac{\hbar}{2\omega_k}} \left(\tilde{e}_k b_k e^{-i\omega_k t} + \tilde{e}^*_{-k} b^*_{-k} e^{+i\omega_k t} \right) \frac{e^{i\vec{k}\cdot\vec{r}}}{\sqrt{V}}$$

This last equation can be rewritten by using the following observations

(1) The summation is over all allowed wave vectors (i.e., all positive and negative wave vector components) so that, *for the second term*, we can replace $-\vec{k} \to \vec{k}$ and $\sum_{-\vec{k}} \to \sum_{\vec{k}}$

(2) For this chapter, assume that the polarization vector is real $\tilde{e}^*_k = \tilde{e}_k$

(3) The angular frequency is positive $\omega_{-\vec{k}} = \omega_{\vec{k}}$

The vector potential becomes

$$\vec{A}(\vec{r},t) = \frac{1}{\sqrt{\varepsilon_o V}} \sum_{\vec{k}} \sqrt{\frac{\hbar}{2\omega_k}} \, \tilde{e}_k \left[\left(b_k e^{-i\omega_k t} \right) e^{i\vec{k}\cdot\vec{r}} + \left(b^*_k e^{+i\omega_k t} \right) e^{-i\vec{k}\cdot\vec{r}} \right] \qquad (3.3.10)$$

or, including the summation over the polarization,

$$\vec{A}(\vec{r},t) = \frac{1}{\sqrt{\varepsilon_o V}} \sum_{\vec{k},s} \sqrt{\frac{\hbar}{2\omega_k}} \, \tilde{e}_{ks} \left[\left(b_{ks} e^{-i\omega_k t} \right) e^{i\vec{k}\cdot\vec{r}} + \left(b^*_{ks} e^{+i\omega_k t} \right) e^{-i\vec{k}\cdot\vec{r}} \right.$$

The equations can be simplified by defining the functions

$$b_{ks}(t) = b_{ks} e^{-i\omega_k t} \qquad \text{and} \qquad b^*_{ks}(t) = b^*_{ks} e^{+i\omega_k t}$$

so that

$$b_{ks} = b_{ks}(0) \qquad \text{and} \qquad b^*_{ks} = b^*_{ks}(0)$$

The coefficients b_{ks} and b^*_{ks} later become the annihilation and creation operators in the quantum theory of light. The annihilation operator removes a photon from a mode characterized by the wave vector $\vec{k}$ and the polarization $\tilde{e}_{ks}$.

The electric and magnetic fields obtain from the relations between the fields and the potentials in a source-free region

$$\vec{E} = -\frac{\partial \vec{A}(\vec{r},t)}{\partial t} \qquad \text{and} \qquad \vec{B} = \nabla \times \vec{A}(\vec{r},t)$$

Therefore,

$$\vec{E} = -\frac{\partial}{\partial t}\vec{A}(\vec{r},t) = \frac{-1}{\sqrt{\varepsilon_o V}}\sum_{\vec{k}}\sqrt{\frac{\hbar}{2\omega_k}}\ \frac{\partial}{\partial t}\left[b_k\, e^{i\vec{k}\cdot\vec{r}-i\omega_k t} + b_k^*\, e^{-i\vec{k}\cdot\vec{r}+i\omega_k t}\right]\tilde{e}_k$$
$$= i\sum_{\vec{k}}\sqrt{\frac{\hbar\omega_k}{2\varepsilon_o V}}\left[b_k\, e^{i\vec{k}\cdot\vec{r}-i\omega_k t} - b_k^*\, e^{-i\vec{k}\cdot\vec{r}+i\omega_k t}\right]\tilde{e}_k \qquad (3.3.11)$$

Similarly, we can calculate the *quantized magnetic field operator* (in the Coulomb gauge)

$$\vec{B} = \nabla\times\vec{A}(\vec{r},t) = \frac{1}{\sqrt{\varepsilon_o V}}\sum_{\vec{k}}\sqrt{\frac{\hbar}{2\omega_k}}\ \nabla\times\left[b_k\, e^{i\vec{k}\cdot\vec{r}-i\omega_k t} + b_k^*\, e^{-i\vec{k}\cdot\vec{r}+i\omega_k t}\right]\tilde{e}_k$$
$$= \frac{1}{\sqrt{\varepsilon_o V}}\sum_{\vec{k}}\sqrt{\frac{\hbar}{2\omega_k}}\ \nabla\left[b_k\, e^{i\vec{k}\cdot\vec{r}-i\omega_k t} + b_k^*\, e^{-i\vec{k}\cdot\vec{r}+i\omega_k t}\right]\times\tilde{e}_k \qquad (3.3.12)$$
$$= i\sum_{\vec{k}}\sqrt{\frac{\hbar}{2\omega_k\varepsilon_o V}}\left(\vec{k}\times\tilde{e}_k\right)\left[b_k\, e^{i\vec{k}\cdot\vec{r}-i\omega_k t} - b_k^*\, e^{-i\vec{k}\cdot\vec{r}+i\omega_k t}\right]$$

Topic 3.3.2: Spatial-Temporal Modes

The previous topic shows the vector potential solution to a wave equation (Equation 3.3.10)

$$\nabla^2\vec{A}_{\vec{k}} - \frac{1}{c^2}\frac{\partial^2\vec{A}_{\vec{k}}}{\partial t^2} = 0$$

using spatial-temporal modes which are traveling plane waves

$$U(x,t) = \frac{e^{i\vec{k}\cdot\vec{r}-i\omega_k t}}{\sqrt{V}}$$

The wave equation can have solutions, which are not traveling waves, depending on the boundary conditions. For example, a perfect no-loss Fabry-Perot cavity has standing sinusoidal solutions; the boundary conditions, in this case, require the fields to be zero at the boundaries. It is important to be able to identify creation and annihilation operators and quantize arbitrary light fields. This topic shows how to write the vector potential in terms of other spatial-temporal modes U(x,t).

$$A(x,t) = \sum_k\sqrt{\frac{\hbar}{2\varepsilon_o\omega_k}}\left[b_k\ U_k(x,t) + b_k^*\ U_k^*(x,t)\right]$$

or as

$$A(x,t) = \sum_k\sqrt{\frac{\hbar}{2\varepsilon_o\omega_k}}\left[b_k\, e^{-i\omega_k t} + b_k^*\, e^{+i\omega_k t}\right]\phi_k(x)$$

where $\phi_k(x)$ is a basis set and

$$b_{ks} = b_{ks}(0) \qquad \text{and} \qquad b_{ks}^* = b_{ks}^*(0)$$

As shown in the next sections, the creation operator $\hat{b}^+$ creates a photon in the mode U_k while the annihilation operator removes a photon from the mode U_k.

The general solution to the wave equation can be found by separating variables and applying the boundary conditions. To solve the wave equation

$$\frac{\partial^2 A(x,t)}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 A(x,t)}{\partial t^2} = 0 \qquad (3.3.13)$$

(where the vector nature of $\vec{A}$ is ignored for simplicity), separate variables according to

$$A_k(x,t) = \phi_k(x)T_k(t) \qquad (3.3.14)$$

Substituting Equation 3.3.14 into the wave equation (3.3.13), separating variables, and taking $-\lambda_k$ as the separation constant (where λ_k>0), gives

$$\frac{1}{\phi_k^2}\frac{\partial^2 \phi_k(x)}{\partial x^2} = -\lambda_k = \frac{1}{c^2}\frac{1}{T_k^2}\frac{\partial^2 T_k(t)}{\partial t^2}$$

The Sturm-Liouville problem for ϕ_k which includes specific boundary conditions provides the set of basis functions $\{\phi_k(x)\}$. The solution of the Sturm-Liouville problem provides the eigenvalues $\lambda_k = k^2$. The solution of the separated time equation

$$\frac{\partial^2 T_k(t)}{\partial t^2} = -\lambda_k c^2 T_k^2$$

is therefore found to be

$$T_k(t) = \sqrt{\frac{\hbar}{2\varepsilon_o \omega_k}}\left[b_k e^{-i\omega_k t} + b_k^* e^{+i\omega_k t}\right]$$

where $\omega_k = ck$. A normalization factor is included for MKS units and $b_{ks} = b_{ks}(0)$ and $b_{ks}^* = b_{ks}^*(0)$. Therefore the general solution of the wave equation is

$$A(x,t) = \sum_k \sqrt{\frac{\hbar}{2\varepsilon_o \omega_k}}\left[b_k e^{-i\omega_k t} + b_k^* e^{+i\omega_k t}\right]\phi_k(x)$$

where $c = \omega_k / k$. For the vector potential to be real (A=A*), the eigenvectors ϕ_k must be real. Therefore, the general spatial-temporal mode is

$$U_k(x,t) = \phi_k(x)e^{-i\omega_k t}$$

so that

$$A(x,t) = \sum_k \sqrt{\frac{\hbar}{2\varepsilon_o \omega_k}}\left[b_k U_k(x,t) + b_k^* U_k^*(x,t)\right]$$

It is perhaps more convenient to write this in the usual notation

$$A(x,t) = \sum_k \sqrt{\frac{\hbar}{2\varepsilon_o \omega_k}}\left[b_k(0) e^{-i\omega_k t} + b_k^*(0) e^{+i\omega_k t}\right]\phi_k(x)$$

The next topic shows that b_k and b_k^* become the annihilation $\hat{b}_k$ and the creation $\hat{b}_k^+$ operators, respectively, in the quantum theory of light. The annihilation operator removes one photon from the mode U_k while the creation operator adds one photon. Because photons are bosons, any number of them can occupy a single state; electrons are Fermions and only one can occupy a given state at a given time.

Including the polarization vector with the basis set, the vector potential

$$\vec{A}(x,t) = \sum_{ks} \sqrt{\frac{\hbar}{2\varepsilon_o \omega_k}}\left[b_{ks}(0) e^{-i\omega_k t} + b_{ks}^*(0) e^{+i\omega_k t}\right]\vec{\phi}_{ks}(x)$$

provides the free-space electric and magnetic fields (in the Coulomb gauge)

$$\vec{E}(x,t) = -\frac{\partial}{\partial t}\vec{A}(x,t) = i\sum_{ks}\sqrt{\frac{\hbar}{2\varepsilon_o\omega_k}}\,\omega_k\left[b_{ks}(0)e^{-i\omega_k t} - b_{ks}^*(0)e^{+i\omega_k t}\right]\vec{\phi}_{ks}(x)$$

and

$$\vec{B}(x,t) = \nabla\times\vec{A}(x,t) = \sum_{ks}\sqrt{\frac{\hbar}{2\varepsilon_o\omega_k}}\left[b_{ks}(0)e^{-i\omega_k t} + b_{ks}^*(0)e^{+i\omega_k t}\right]\nabla\times\vec{\phi}_{ks}(x)$$

Section 3.4: The Quantized Free-Field Hamiltonian

The typical method for transforming the classical Hamiltonian into the corresponding quantum mechanical one consists of replacing the classical dynamical variables and operators and then requiring these operators to satisfy commutation rules. The next chapter begins to explore the Hilbert space upon which the operators act (the simplest being Fock space).

Topic 3.4.1: The Classical Free-Field Hamiltonian

The divergence of the Poynting vector leads to an expression for the electromagnetic power flowing into/out of a volume (same result obtains using the Lagrangian for the electromagnetic fields given in the appendices). The energy density in free-space (MKS units) is identified as

$$H_c = \frac{\varepsilon_o}{2}\vec{E}\cdot\vec{E} + \frac{1}{2\mu_o}\vec{B}\cdot\vec{B}$$

so that the total energy in a volume V is

$$H_c = \int_V dV\left(\frac{\varepsilon_o}{2}\vec{E}\cdot\vec{E} + \frac{1}{2\mu_o}\vec{B}\cdot\vec{B}\right) \tag{3.4.1}$$

where the subscript "c" refers to the classical case. Section 7.2 shows that the vector potential

$$\vec{A}(\vec{r},t) = \frac{1}{\sqrt{\varepsilon_o V}}\sum_{\vec{k}s}\sqrt{\frac{\hbar}{2\omega_k}}\,\tilde{e}_{ks}\left[b_{ks}(t)\,e^{i\vec{k}\cdot\vec{r}} + b_{ks}^*(t)\,e^{-i\vec{k}\cdot\vec{r}}\right]$$

for a free-space traveling wave leads to the classical electric field

$$\vec{E} = -\frac{\partial}{\partial t}\vec{A}(\vec{r},t) = \frac{+i}{\sqrt{\varepsilon_o V}}\sum_{\vec{k}s}\sqrt{\frac{\hbar\omega_k}{2}}\left[b_{ks}(t)e^{i\vec{k}\cdot\vec{r}} - b_{ks}^*(t)e^{-i\vec{k}\cdot\vec{r}}\right]\tilde{e}_{ks} \tag{3.4.2}$$

and to the classical magnetic field

$$\vec{B} = \nabla\times\vec{A}(\vec{r},t) = \frac{+i}{\sqrt{\varepsilon_o V}}\sum_{\vec{k}s}\sqrt{\frac{\hbar}{2\omega_k}}\left(\vec{k}\times\tilde{e}_{ks}\right)\left[b_{ks}(t)e^{i\vec{k}\cdot\vec{r}} - b_{ks}^*(t)e^{-i\vec{k}\cdot\vec{r}}\right] \tag{3.4.3}$$

where the index "s" refers to the polarization of the mode and

$$b_{ks}(t) = b_{ks}(0)e^{-i\omega_k t} = b_{ks}e^{-i\omega_k t} \qquad b_{ks}^*(t) = b_{ks}^*(0)e^{i\omega_k t} = b_{ks}^*e^{i\omega_k t}$$

Keep in mind, that

$$\left\{u_{\vec{k}}(\vec{r}) = \frac{e^{i\vec{k}\cdot\vec{r}}}{\sqrt{V}}\right\}$$

consists of discrete basis vectors with the orthonormality relation of

$$\delta_{\vec{k}\vec{K}} = \left\langle u_{\vec{k}}(\vec{r})\middle| u_{\vec{K}}(\vec{r})\right\rangle = \int_V dV\, u_{\vec{k}}^*(\vec{r})\,u_{\vec{K}}(\vec{r}) = \int_V dV\,\frac{e^{i(\vec{K}-\vec{k})\vec{r}}}{V} \tag{3.4.4}$$

The classical Hamiltonian can be written in terms of the Fourier amplitudes by substituting for the electric and magnetic fields in Equation 3.4.1. The term involving the magnetic field is slightly more involved to calculate and that's the one we do here; the

term with the electric field is handled similarly. Drop the time-dependence of the Fourier coefficients b and b^* to make the notation more compact until Equation 3.4.7.

$$\int_V dV \frac{1}{2\mu_o}\vec{B}\cdot\vec{B} = \frac{-\hbar}{4\mu_o\varepsilon_o V}\int_V dV \sum_{\vec{K}S}\sum_{\vec{k}s}\frac{1}{\sqrt{\omega_k\omega_K}}\left(\vec{k}\times\tilde{e}_{ks}\right)\left(\vec{K}\times\tilde{e}_{KS}\right)\left[b_{ks}\,e^{i\vec{k}\cdot\vec{r}} - b^*_{ks}\,e^{-i\vec{k}\cdot\vec{r}}\right]\left[b_{KS}\,e^{i\vec{K}\cdot\vec{r}} - b^*_{KS}\,e^{-i\vec{K}\cdot\vec{r}}\right]$$

$$= \frac{-\hbar}{4\mu_o\varepsilon_o V}\int_V dV \sum_{\vec{K}S}\sum_{\vec{k}s}\frac{1}{\sqrt{\omega_k\omega_K}}\left(\vec{k}\times\tilde{e}_{ks}\right)\left(\vec{K}\times\tilde{e}_{KS}\right)\left[b_{ks}b_{KS}\,e^{i(\vec{k}+\vec{K})\vec{r}} + b^*_{ks}\,b^*_{KS}e^{-i(\vec{k}+\vec{K})\vec{r}} + \right.$$

$$\left. - b_{ks}b^*_{KS}\,e^{i(\vec{k}-\vec{K})\vec{r}} - b^*_{ks}\,b_{KS}e^{-i(\vec{k}-\vec{K})\vec{r}}\right]$$

Using the orthonormality relation in Equation 3.4.4, this equation simplifies to

$$\int_V dV \frac{1}{2\mu_o}\vec{B}\cdot\vec{B} = \frac{-\hbar}{4\mu_o\varepsilon_o}\sum_{\vec{k}sS}\frac{1}{\omega_k}\left[\left(\vec{k}\times\tilde{e}_{ks}\right)\left(-\vec{k}\times\tilde{e}_{-kS}\right)b_{ks}b_{-\vec{k}S} + \left(\vec{k}\times\tilde{e}_{ks}\right)\left(-\vec{k}\times\tilde{e}_{-kS}\right)b^*_{ks}\,b^*_{-kS} + \right.$$

$$\left. - \left(\vec{k}\times\tilde{e}_{ks}\right)\left(\vec{k}\times\tilde{e}_{kS}\right)b_{ks}b^*_{kS} - \left(\vec{k}\times\tilde{e}_{ks}\right)\left(\vec{k}\times\tilde{e}_{kS}\right)b^*_{ks}\,b_{kS}\right]$$

where we have used the fact that $\omega_k = \omega_{-k}$. Next using the general vector relation

$$\left(\vec{A}\times\vec{B}\right)\cdot\left(\vec{C}\times\vec{D}\right) = \left(\vec{A}\cdot\vec{C}\right)\left(\vec{B}\cdot\vec{D}\right) - \left(\vec{A}\cdot\vec{D}\right)\left(\vec{B}\cdot\vec{C}\right) \qquad (3.4.5)$$

and the fact that the polarization vectors satisfy and orthonormality relation

$$\tilde{e}_{ks}\cdot\tilde{e}_{KS} = \delta_{sS} \qquad (3.4.6)$$

since different polarizations are orthogonal regardless of the direction of motion. We find

$$\left(\vec{k}\times\tilde{e}_{ks}\right)\left(\vec{k}\times\tilde{e}_{kS}\right) = k^2\delta_{sS} \qquad \text{and} \qquad \left(\vec{k}\times\tilde{e}_{ks}\right)\left(-\vec{k}\times\tilde{e}_{-kS}\right) = -k^2\delta_{sS}$$

The integral over the magnetic field becomes

$$\int_V dV \frac{1}{2\mu_o}\vec{B}\cdot\vec{B} = \frac{-\hbar}{4\mu_o\varepsilon_o}\sum_{\vec{k}s}\frac{1}{\omega_k}\left[-k^2\,b_{ks}b_{-\vec{k}s} - k^2\,b^*_{ks}\,b^*_{-ks} - k^2 b_{ks}b^*_{ks} - k^2 b^*_{ks}\,b_{ks}\right]$$

Making the substitution

$$\frac{1}{\mu_o\varepsilon_o} = c^2 = \left(\frac{\omega_k}{k}\right)^2$$

the result for the energy residing in the magnetic field is

$$\int_V dV \frac{1}{2\mu_o}\vec{B}\cdot\vec{B} = \frac{-\hbar}{4}\sum_{\vec{k}s}\omega_k\left[-b_{ks}b_{-\vec{k}s} - b^*_{ks}\,b^*_{-ks} - b_{ks}b^*_{ks} - b^*_{ks}\,b_{ks}\right]$$

In a similar manner, but with a lot less trouble, the expression for the integral over the electric field is

$$\int_V dV \frac{\varepsilon_o}{2}\vec{E}\cdot\vec{E} = \frac{-\hbar}{4}\sum_{\vec{k}s}\omega_k\left[+b_{ks}b_{-\vec{k}s} + b^*_{ks}\,b^*_{-ks} - b_{ks}b^*_{ks} - b^*_{ks}\,b_{ks}\right]$$

Therefore, the classical Hamiltonian in the Equation 3.4.1 becomes

$$H_c = \int_V dV\left(\frac{\varepsilon_o}{2}\vec{E}\cdot\vec{E} + \frac{1}{2\mu_o}\vec{B}\cdot\vec{B}\right) = \frac{1}{2}\sum_{\vec{k}s}\hbar\omega_k\left[b^*_{\vec{k}s}(t)b_{\vec{k}s}(t) + b_{\vec{k}s}(t)b^*_{\vec{k}s}(t)\right] \qquad (3.4.7)$$

where we have been careful not to commute the conjugate variables since they will become creation and annihilation operators which do not commute.

Topic 3.4.2: The Quantum Mechanical Free-Field Hamiltonian

Replacing the classical fields with operators quantizes the classical Hamiltonian density

$$H_c = \frac{\varepsilon_o}{2}\hat{E}^2 + \frac{1}{2\mu_o}\hat{B}^2$$

Replacing the Fourier amplitudes with the corresponding creation and annihilation operators quantizes the classical Hamiltonian in Equation 3.4.7

$$\hat{H} = \frac{1}{2}\sum_{\vec{k}s}\hbar\omega_k\left[\hat{b}^+_{\vec{k}s}(t)\hat{b}_{\vec{k}s}(t) + \hat{b}_{\vec{k}s}(t)\hat{b}^+_{\vec{k}s}(t)\right] = \frac{1}{2}\sum_{\vec{k}s}\hbar\omega_k\left[\hat{b}^+_{\vec{k}s}\,\hat{b}_{\vec{k}s} + \hat{b}_{\vec{k}s}\,\hat{b}^+_{\vec{k}s}\right.$$

The creation and annihilation operators depend on time according to

$$\hat{b}_{ks}(t) = \hat{b}_{ks}(0)e^{-i\omega_k t} = \hat{b}_{ks}\,e^{-i\omega_k t} \qquad \text{and} \qquad \hat{b}^+_{ks}(t) = \hat{b}^+_{ks}(0)e^{+i\omega_k t} = \hat{b}^+_{ks}\,e^{+i\omega_k t}$$

(the units of H_c is energy are "per unit volume" and the units of H are energy). Notice that the time-dependence in the free-field Hamiltonian cancels. The time-dependent annihilation and creation operators are Heisenberg operators (refer to the next example) with required equal-time commutation relations

$$\left[\hat{b}_{\vec{k}s}(t), \hat{b}^+_{\vec{K}S}(t)\right] = \delta_{\vec{k}\vec{K}}\delta_{sS}$$

$$\left[\hat{b}_{\vec{k}s}(t), \hat{b}_{\vec{K}S}(t)\right] = 0 = \left[\hat{b}^+_{\vec{k}s}(t), \hat{b}^+_{\vec{K}S}(t)\right]$$

hold for all times including t=0. Many expression in quantum optics require "normally ordered" creation and annihilation operators where the creation operators are left of the annihilation operators. Therefore the first commutation relation

$$\hat{b}_{\vec{k}s}\hat{b}^+_{\vec{k}s} = \hat{b}^+_{\vec{k}s}\hat{b}_{\vec{k}s} + 1$$

is used to change the second term in Hamiltonian $\hat{H}$ so that

$$\hat{H} = \sum_{\vec{k}s}\hbar\omega_k\left(\hat{b}^+_{\vec{k}s}\hat{b}_{\vec{k}s} + \frac{1}{2}\right) = \sum_{\vec{k}s}\hbar\omega_k\left(\hat{N}_{\vec{k}s} + \frac{1}{2}\right) \qquad (3.4.8)$$

Notice that the Hamiltonian for light is similar to the Hamiltonian for the harmonic oscillator. The summation occurs because light has many modes as opposed to the single mode for the massive-particle harmonic oscillator discussed in Chapter 2. Therefore the electromagnetic field appears as an ensemble of *independent* harmonic oscillators. The oscillators are "independent " because there are no cross-terms between modes in Equation 3.4.8. The number operator $\hat{N}_{\vec{k}s}$ provides the number of photons in a particular mode specified by the wave vector $\vec{k}$ and polarization "s

Equation 3.4.8

$$\hat{H} = \sum_{\vec{k}s}\hbar\omega_k\left(\hat{N}_{\vec{k}s} + \frac{1}{2}\right)$$

contains a summation over the frequency for all of the possible modes, namely,

$$\frac{1}{2}\sum_{\vec{k}s}\hbar\omega_k \qquad (3.4.9)$$

The allowed frequencies can be infinitely large. This energy is stored as a fluctuating electric field in the vacuum state as discussed in Section 7.2. The summation in Equation 3.4.9 becomes infinite even for the finite volume V of integration initially used to

calculate the energy! Physically, this implies a very large energy stored in the vacuum! In some cases, it is possible to ignore the divergent term. For example, the rate of change of an operator which involves the commutator with Hamiltonian is insensitive to the infinity.

Example: Calculate the time-dependence of the creation operator $b_K^+(t)$ in the Heisenberg picture for the free-fields.

Solution: The simplest method is to calculate the commutator of the creation operator with the Hamiltonian. Chapter 6 Section 4 Topic 3 shows that the rate of change of the Heisenberg operator is given by

$$\frac{d\hat{b}_K^+(t)}{dt} = \frac{i}{\hbar}\left[\hat{H}, \hat{b}_K^+(t)\right]$$

Substituting the Hamiltonian

$$\hat{H} = \sum_{\bar{k}} \hbar\omega_k \left(\hat{b}_{\bar{k}}^+(t)\hat{b}_{\bar{k}}(t) + \frac{1}{2} \right)$$

gives

$$\frac{d\hat{b}_{\bar{K}}^+(t)}{dt} = \frac{i}{\hbar}\left[\sum_{\bar{k}} \hbar\omega_k \left(\hat{b}_{\bar{k}}^+(t)\hat{b}_{\bar{k}}(t) + \frac{1}{2} \right), \hat{b}_{\bar{K}}^+(t) \right] = \frac{i}{\hbar} \sum_{\bar{k}} \hbar\omega_k \left\{ \left[\hat{b}_{\bar{k}}^+(t)\hat{b}_{\bar{k}}(t), \hat{b}_{\bar{K}}^+(t) \right] + \left[\frac{1}{2}, \hat{b}_{\bar{K}}^+(t) \right] \right\}$$

The infinite vacuum-energy term gives the last commutator which is zero; consequently, the infinite term does not affect the calculated value of $\hat{b}^+$. Using commutation rules, we can evaluate

$$\left[\hat{b}_{\bar{k}}^+\hat{b}_{\bar{k}}, \hat{b}_{\bar{K}}^+ \right] = \left[\hat{b}_{\bar{k}}^+, \hat{b}_{\bar{K}}^+ \right]\hat{b}_{\bar{k}} + \hat{b}_{\bar{k}}^+\left[\hat{b}_{\bar{k}}, \hat{b}_{\bar{K}}^+ \right] = 0 + \hat{b}_{\bar{k}}^+ \delta_{\bar{k}\bar{K}}$$

Substituting this into the previous expression gives

$$\frac{d\hat{b}_{\bar{K}}^+(t)}{dt} = \frac{i}{\hbar} \sum_{\bar{k}} \hbar\omega_k \left[\hat{b}_{\bar{k}}^+(t)\hat{b}_{\bar{k}}(t), \hat{b}_{\bar{K}}^+(t) \right] = i\omega_K \hat{b}_{\bar{K}}^+(t)$$

This is a simple differential equation with a solution that agrees with our previous results

$$\hat{b}_{\bar{K}}^+(t) = \hat{b}_{\bar{K}}^+(0)e^{i\omega_K t}.$$

The time dependence of the creation and annihilation operators depend on the Hamiltonian. The results for $\hat{b}^+$ have been calculated using the free-field Hamiltonian. The creation and annihilation operators have a different time-dependence if the interaction between the fields and matter is included. The reason for this is that the matter can itself produce electromagnetic fields, which necessarily change the field operators. See for example the section on the Jayne-Cumming's model.

Topic 3.4.3: The EM Hamiltonian with Quadrature Operators

This section introduces the notion of "position and momentum" operators for the electromagnetic field. These operators do not refer to photon position and momentum. Instead, they refer to the "out-of-phase" and "in-phase" quadrature components of the electric field; this is discussed in greater detail in subsequent chapters. These operators are important since, besides squeezing noise from the amplitude or phase, noise can be squeezed from the quadrature components. In addition, a Schrodinger equation can be set up based on the "coordinate representation" of the Hamiltonian.

As previously mentioned, Equation 3.4.8

$$\hat{H} = \sum_{\vec{k}s} \hbar\omega_k \left(\hat{b}^+_{\vec{k}s} \hat{b}_{\vec{k}s} + \frac{1}{2} \right) = \sum_{\vec{k}s} \hbar\omega_k \left(\hat{N}_{\vec{k}s} + \frac{1}{2} \right)$$

has the same form as that for a collection of independent harmonic oscillators. Similar to the harmonic oscillator, the creation and annihilation operators can be related to "position" $\hat{q}_k$ and "momentum" $\hat{p}_k$ operators according to

$$\hat{b}_{ks}(t) = \frac{\omega_k \hat{q}_{ks}(t)}{\sqrt{2\hbar\omega_k}} + \frac{i\hat{p}_{ks}(t)}{\sqrt{2\hbar\omega_k}} \qquad (3.4.10)$$

and, by taking the adjoint,

$$\hat{b}^+_{ks}(t) = \frac{\omega_k \hat{q}_{ks}(t)}{\sqrt{2\hbar\omega_k}} - \frac{i\hat{p}_{ks}(t)}{\sqrt{2\hbar\omega_k}} \qquad (3.4.11)$$

where $\hat{q}_{ks}(t)$ and $\hat{p}_{ks}(t)$ are taken to be Hermitian operators. Equations 3.4.10 and 3.4.11 hold for t=0 with the definitions

$$q_{ks} = q_{ks}(0) \quad \text{and} \quad p_{ks} = p_{ks}(0)$$

The subscripts "k" and "s" label the wavelength and polarization modes, respectively. The subscript "s" is usually suppressed for simplicity. *The position* $\hat{q}_k$ *and momentum* $\hat{p}_k$ *are not related to the spatial position* $\vec{r}$ *or the photon momentum* $\hbar\vec{k}$. Solving Equations 3.4.10 and 3.4.11 for the *position* $\hat{q}_k$ *and momentum* $\hat{p}_k$ provides

$$\hat{q}_k(t) = \sqrt{\frac{\hbar}{2\omega_k}} \left(\hat{b}_k(t) + \hat{b}^+_k(t) \right)$$

$$\hat{p}_k(t) = -i\sqrt{\frac{\hbar\omega_k}{2}} \left(\hat{b}_k(t) - \hat{b}^+_k(t) \right)$$

which is similar to those for the harmonic oscillator except here, there is no mass.

The commutation relations for the creation and annihilation operators provide the commutation relations between the position and momentum operators as follows

$$\left[\hat{q}_i(t), \hat{q}_j(t)\right] = 0 = \left[\hat{p}_i(t), \hat{p}_j(t)\right] \qquad \text{for all modes i,j}$$

$$\left[\hat{q}_i(t), \hat{p}_j(t)\right] = i\hbar\delta_{ij}$$

which hold for all times t including t=0.

Example: Prove the last commutation relation

Solution:

$$\left[\hat{q}_i,\hat{p}_j\right]=\left[\sqrt{\frac{\hbar}{2\omega_i}}\left(\hat{b}_i+\hat{b}_i^+\right),-i\sqrt{\frac{\hbar\omega_i}{2}}\left(\hat{b}_j-\hat{b}_j^+\right)\right]$$

$$=-i\sqrt{\frac{\hbar}{2\omega_i}}\sqrt{\frac{\hbar\omega_j}{2}}\left[\hat{b}_i+\hat{b}_i^+,\hat{b}_j-\hat{b}_j^+\right]$$

$$=-i\frac{\hbar}{2}\sqrt{\frac{\omega_j}{\omega_i}}\left\{\left[\hat{b}_i,-\hat{b}_j^+\right]+\left[\hat{b}_i^+,\hat{b}_j\right]\right\}=-i\frac{\hbar}{2}\left\{-\delta_{ij}+-\delta_{ij}\right\}$$

$$=i\hbar\delta_{ij}$$

The Hamiltonian and the fields can be written in terms of the position and momentum operators. Starting with Hamiltonian

$$\hat{H}=\sum_{\bar{k}s}\hbar\omega_k\left(\hat{b}_{\bar{k}s}^+\hat{b}_{\bar{k}s}+\frac{1}{2}\right)=\sum_{\bar{k}s}\hbar\omega_k\left(\hat{N}_{\bar{k}s}+\frac{1}{2}\right)$$

Neglecting the polarization index and substituting the position and momentum expressions for the creation and annhilation operators in the Hamiltonian provides

$$\hat{H}=\sum_{\bar{k}}\hbar\omega_k\left[\left(\frac{\omega_k\,\hat{q}_k}{\sqrt{2\hbar\omega_k}}+\frac{i\,\hat{p}_k}{\sqrt{2\hbar\omega_k}}\right)\left(\frac{\omega_k\,\hat{q}_k}{\sqrt{2\hbar\omega_k}}-\frac{i\,\hat{p}_k}{\sqrt{2\hbar\omega_k}}\right)+\frac{1}{2}\right]$$

Multiplying out the terms and taking care not to commute non-commuting operators

$$\hat{H}=\sum_{\bar{k}}\hbar\omega_k\left[\frac{\omega_k^2\,\hat{q}_k^2}{2\hbar\omega_k}+\frac{\hat{p}_k^2}{2\hbar\omega_k}-i\frac{\omega_k}{2\hbar\omega_k}\left(\hat{q}_k\hat{p}_k-\hat{p}_k\hat{q}_k\right)+\frac{1}{2}\right]$$

Using the commutation relation $\left[\hat{q}_a,\hat{p}_b\right]=i\hbar\delta_{ab}$ and then simplifying gives

$$\hat{H}=\sum_{\bar{k}}\left(\frac{\hat{p}_{\bar{k}}^2}{2}+\frac{\omega_k^2}{2}q_{\bar{k}}^2\right) \tag{3.4.12}$$

The Hamiltonian is composed of a sum of Hamiltonians for a collection of independent harmonic oscillators.

A Schrodinger equation can be developed for the electromagnetic field. The "coordinate representation" of the "position and momentum" operators provides

$$\hat{q}_k\rightarrow q_k \qquad \hat{p}_k\rightarrow\frac{\hbar}{i}\frac{\partial}{\partial q_k}$$

Here, the coordinate q_k is thought of as an electric field amplitude and not as a photon position. The coordinated representation of the position and momentum operator can be substituted into the Hamiltonian to obtain the coordinate representation of the Schrodinger equation

$$\sum_{\bar{k}}\left(-\frac{\hbar^2}{2}\frac{\partial^2}{\partial q_{\bar{k}}^2}+\frac{\omega_k^2}{2}q_{\bar{k}}^2\right)\Psi(q_1,q_2\ldots,t)=i\hbar\frac{\partial}{\partial t}\Psi(q_1,q_2\ldots,t)$$

The solutions of this wave equation are expected to be similar to that for the harmonic oscillator. We expect decaying exponentials multiplied by Hermite polynomials. The next couple of chapters discusses the solutions.

Section 3.5: Quantized Vector Potential and Free-Fields

The previous section shows how the Hamiltonian is quantized by substituting operators $\hat{E},\hat{B}$ for the electric and magnetic fields in the classical expression for the energy. The procedure is equivalent to substituting creation and annihilation operators for the Fourier amplitudes. The creation and annihilation operators are required to satisfy certain commutation relations. Similarly, the Hamiltonian can be written in terms of generalized position and momentum operators that describe the field amplitudes. Likewise, the generalized position and momentum operators satisfy commutation relations. These commutation relations are responsible for Heisenberg uncertainty relations that exist between conjugate variables. Trying to simultaneously measure non-commuting operators necessarily results in a distribution of measured values. For coherent states, the distribution is described by Poisson statistics. The Wigner distributions are the closest quantum analogue of the classical probability distribution.

Topic 3.5.1: The Vector Potential

This section explicitly discusses the conversion of the classical vector potential into an operator. Quantize fields operate on a Hilbert space the simplest of which is the Fock space.

From the previous section, the Fourier components of the *classical* vector potential are

$$b_k(t) = b_k\, e^{-i\omega_k t} = b_k(0)e^{-i\omega_k t}$$
$$b_k^*(t) = b_k^*\, e^{+i\omega_k t} = b_k^*(0)e^{+i\omega_k t}$$

or, equivalently said, "b" *is the amplitude* of a given optical mode "k". The derivation assumes a source-free region of space (i.e., no interaction potential). Replacing the amplitudes with operators according to the prescription

$$b_k \to \hat{b}_k \quad \text{and} \quad b_k^* \to \hat{b}_k^+$$

quantizes the vector potential. The time dependent operators

$$\begin{aligned} \hat{b}_k(t) &= \hat{b}_k\, e^{-i\omega_k t} = \hat{b}_k(0)e^{-i\omega_k t} \\ \hat{b}_k^+(t) &= \hat{b}_k^+\, e^{+i\omega_k t} = \hat{b}_k^+(0)e^{+i\omega_k t} \end{aligned} \tag{3.5.1}$$

must satisfy the commutation relations

$$\left[\hat{b}_\xi(t), b_\eta(t)\right] = 0 = \left[\hat{b}_\xi^+(t), \hat{b}_\eta^+(t)\right] \qquad \text{for all } \xi,\eta \tag{3.5.2}$$

and

$$\left[\hat{b}_\xi(t), \hat{b}_\eta^+(t)\right] = \delta_{\xi\eta} \tag{3.5.3}$$

Notice the commutators are evaluated at a single time; the commutators are "equal time" commutators. When the creation and annihilation operators are given in the Heisenberg representation, the commutation relations can be different at unequal times: operators that once commuted at equal time might not commute at different times. Equations 3.5.2 and 3.5.3 show that two modes with $\xi \neq \eta$ are considered to be independent so that creating a photon in state ξ is independent of annihilating a photon in the state η. However, as previously discussed, a Heisenberg uncertainty relation should be expected for the mode $\xi=\eta$ owing to the last commutation relation. Notice that the boundary conditions lead to

discrete values for the wave vector $\vec{k}$ which in turn lead to the Kronecker Delta function (as opposed to the Dirac Delta function). As an important note, equation 3.5.1 defines the Heisenberg representation for the annihilation and creation operators *in free space* (i.e., there is no interaction potential between the fields and matter). The form of Equation 3.5.1 necessarily changes when the fields are allowed to interact with matter since the matter can produce additional fields or absorb energy from the fields (in which case, the amplitudes must change).

The classical vector potential

$$\vec{A}(\vec{r},t)=\frac{1}{\sqrt{\varepsilon_o V}}\sum_{\vec{k}}\sqrt{\frac{\hbar}{2\omega_k}}\,\tilde{e}_k\left[\left(b_k e^{-i\omega_k t}\right)e^{i\vec{k}\cdot\vec{r}}+\left(b_k^* e^{+i\omega_k t}\right)e^{-i\vec{k}\cdot\vec{r}}\right]$$

becomes the operator

$$\hat{A}(\vec{r},t)=\frac{1}{\sqrt{\varepsilon_o V}}\sum_{\vec{k}}\sqrt{\frac{\hbar}{2\omega_k}}\,\tilde{e}_k\left[\hat{b}_{\vec{k}}(t)e^{i\vec{k}\cdot\vec{r}}+\hat{b}_{\vec{k}}^{+}(t)e^{-i\vec{k}\cdot\vec{r}}\right] \qquad (3.5.4)$$

A comment on the structure of quantum theory is in order at this point. The operators in the quantized field theory contained all the possible creation and annihilation operators (i.e., modes). In some sense, the field (and the Hamiltonian) must contain all of the possibilities that can physically occur. The Fock states contain the specifics of the system. If only two modes are occupied, for example, then the Fock state will contain only two nonzero entries.

Topic 3.5.2: Electric and Magnetic Fields

The electromagnetic fields are the "conceptual work horses" for quantum optics. The fields are written in terms of the creation/annihilation operators (generalized amplitudes) or equivalently interms of the generalized position/momentum operators. The next chapter shows that the Fock states are eigenvectors of the number operator and coherent states are the eigenvectors of the annihilation operator.

Changing the Fourier amplitudes into operators quantizes the electromagnetic fields. These operators are thought of as providing the amplitude of the wave. The picture is clarified by specifying a Hilbert space (in particular for coherent states). The previous section discusses the quantization of the vector potential for an electromagnetic wave.

The *quantized electric field operator* (in the Coulomb gauge) is found by differentiating the vector potential with respect to time. The vector potential is

$$\hat{A}(\vec{r},t)=\frac{1}{\sqrt{\varepsilon_o V}}\sum_{\vec{k}}\sqrt{\frac{\hbar}{2\omega_k}}\left[\hat{b}_{\vec{k}}(0)e^{i\vec{k}\cdot\vec{r}-i\omega_k t}+\hat{b}_{\vec{k}}^{+}(0)e^{-i\vec{k}\cdot\vec{r}+i\omega_k t}\right]\tilde{e}_k$$

Differentiating the vector potential with respect to time gives

$$\begin{aligned}\hat{E}=-\frac{\partial}{\partial t}\hat{A}(\vec{r},t)&=\frac{-1}{\sqrt{\varepsilon_o V}}\sum_{\vec{k}}\sqrt{\frac{\hbar}{2\omega_k}}\,\frac{\partial}{\partial t}\left[\hat{b}_k(0)e^{i\vec{k}\cdot\vec{r}-i\omega_k t}+\hat{b}_k^{+}(0)e^{-i\vec{k}\cdot\vec{r}+i\omega_k t}\right]\tilde{e}_k\\&=+i\sum_{\vec{k}}\sqrt{\frac{\hbar\omega_k}{2\varepsilon_o V}}\left[\hat{b}_k(0)e^{i\vec{k}\cdot\vec{r}-i\omega_k t}-\hat{b}_k^{+}(0)e^{-i\vec{k}\cdot\vec{r}+i\omega_k t}\right]\tilde{e}_k\end{aligned} \qquad (3.5.5)$$

Similarly, the *quantized magnetic field operator* (in the Coulomb gauge) provides

$$\hat{B} = \nabla \times \hat{A}(\vec{r},t) = \frac{1}{\sqrt{\varepsilon_o V}} \sum_{\vec{k}} \sqrt{\frac{\hbar}{2\omega_k}} \; \nabla \times \left[\hat{b}_k(0) e^{i\vec{k}\cdot\vec{r} - i\omega_k t} + \hat{b}_k^+(0) e^{-i\vec{k}\cdot\vec{r} + i\omega_k t} \right] \vec{e}_k$$

$$= \frac{1}{\sqrt{\varepsilon_o V}} \sum_{\vec{k}} \sqrt{\frac{\hbar}{2\omega_k}} \; \nabla \left[\hat{b}_k(0) e^{i\vec{k}\cdot\vec{r} - i\omega_k t} + \hat{b}_k^+(0) e^{-i\vec{k}\cdot\vec{r} + i\omega_k t} \right] \times \vec{e}_k \qquad (3.5.6)$$

$$= +i \sum_{\vec{k}} \sqrt{\frac{\hbar}{2\omega_k \varepsilon_o V}} \left(\vec{k} \times \vec{e}_k \right) \left[\hat{b}_k(0) e^{i\vec{k}\cdot\vec{r} - i\omega_k t} - \hat{b}_k^+(0) e^{-i\vec{k}\cdot\vec{r} + i\omega_k t} \right]$$

Again, the above quantized electric and magnetic fields are for free-space where there is no interaction between the fields and charges. For example, we expect the fields to increase if they interact with atoms that produce stimulated emission. Equation's 3.5.5 and 3.5.6 do not contain a factor that can account for this increase. In this text, the *time independent* creation and annihilation operators are denoted by "b^+" and "b" respectively. For Equations 3.5.5 and 3.5.6, $b^+ = b^+(0)$ and $b = b(0)$.

As mentioned in the introductory material, light is quantized and that the energy is parceled in photons. This means that the fields and Hamiltonian are also quantized. The reader is aware of the typical conceptual problems with trying to picture light as both particle and wave. However, it is interesting to note that both pictures are combined in the quantized fields as shown for example by Equation 3.5.5. The traveling wave portion represented by $e^{ikx - i\omega t}$ represents the wave nature of light whereas the creation and annihilation operators represent the particle nature of light. A thorough treatment of field quantization shows that all particles (not just electrons and photons) have similar equations with both wave and particle terms. And by the way, similar to Equations 3.5.5 and 3.5.6, the traveling wave portion can be replaced by other wave functions such as the sine and cosine for the Fabry Perot cavity. Once again, the electric field operators are seen to contain the annihilation and creation operators for all of the possible modes of the system. It is up to the wave function (i.e., Fock states) to describe the actual physical system (i.e., how many photons and what modes they occupy).

Topic 3.5.3: Other Basis Sets

The previous section shows that the vector potential can be written in other basis sets besides the traveling waves. If the set $\{\phi_n(x)\}$ forms a basis set that satisfies the boundary conditions then the vector potential that satisfies the wave equation can be written as

$$\hat{A}(x,t) = \sum_{ks} \sqrt{\frac{\hbar}{2\varepsilon_o \omega_k}} \left[\hat{b}_{ks}(0) e^{-i\omega_k t} + \hat{b}_{ks}^+(0) e^{+i\omega_k t} \right] \vec{\phi}_{ks}(x)$$

where the polarization vector has been grouped with the basis functions and the creation and annihilation operators replace the Fourier expansion coefficients.
As a result, the free-space electric and magnetic fields (in the Coulomb gauge) can be written as

$$\hat{E}(x,t) = -\frac{\partial}{\partial t} \hat{A}(x,t) = i \sum_{ks} \sqrt{\frac{\hbar\omega_k}{2\varepsilon_o}} \left[\hat{b}_{ks}(0) e^{-i\omega_k t} - \hat{b}_{ks}^+(0) e^{+i\omega_k t} \right] \vec{\phi}_{ks}(x)$$

and

$$\hat{B}(x,t) = \nabla \times \hat{A}(x,t) = \sum_{ks} \sqrt{\frac{\hbar}{2\varepsilon_o \omega_k}} \left[\hat{b}_{ks}(0) e^{-i\omega_k t} + \hat{b}_{ks}^{+}(0) e^{+i\omega_k t} \right] \nabla \times \bar{\phi}_{ks}(x)$$

Example: Find the quantized electric field for the perfect Fabry-Perot cavity with the left mirror at z=0 and the right mirror at z=L.

Solution: The standing wave modes are

$$\bar{\phi}_n(z) = \hat{x} \sqrt{\frac{2}{L}} \sin(k_n z) \qquad \text{where} \quad k_n = \frac{\pi n}{L} \qquad n=1,2,3...$$

The electric field is therefore given by

$$\hat{E} = -\frac{\partial \hat{A}}{\partial t} = i \sum_k \sqrt{\frac{\hbar \omega_k}{2\varepsilon_o}} \left[\hat{b}_k(0) e^{-i\omega_k t} - \hat{b}_k^{+}(0) e^{+i\omega_k t} \right] \bar{\phi}_k(z)$$

$$= i \hat{x} \sum_k \sqrt{\frac{\hbar \omega_k}{\varepsilon_o L}} \left[\hat{b}_k(0) e^{-i\omega_k t} - \hat{b}_k^{+}(0) e^{+i\omega_k t} \right] \sin(kz)$$

where the sum is over the allowed values of "k".

Section 3.6: Fields with Quadrature Operators

Recall from a previous section that the creation and annihilation operators are related to the position $\hat{q}_k$ and momentum $\hat{p}_k$ operators according to

$$\hat{b}_k = \frac{\omega_k \hat{q}_k}{\sqrt{2\hbar\omega_k}} + \frac{i\hat{p}_k}{\sqrt{2\hbar\omega_k}} \tag{3.6.1}$$

$$\hat{b}_k^+ = \frac{\omega_k \hat{q}_k}{\sqrt{2\hbar\omega_k}} - \frac{i\hat{p}_k}{\sqrt{2\hbar\omega_k}} \tag{3.6.2}$$

where $\hat{q}_k$ and $\hat{p}_k$ are Hermitian operators. The subscripts "k" label the modes. The definitions in Equations 3.6.1 and 3.6.2 actually depend on the Hamiltonian for the optical fields being similar to the harmonic oscillator. *The position $\hat{q}_k$ and momentum $\hat{p}_k$ are not related to the spatial position $\vec{r}$ nor are they related to the photon momentum $\hbar\vec{k}$.* Previous equations indicate that these operators are related to the amplitude of the electric and magnetic fields.

The position and momentum operators are related to the fields as can be seen by substituting Equations 3.6.1 and 3.6.2 into the expressions for the fields.

$$\vec{A}(\vec{r},t) = \frac{1}{\sqrt{\varepsilon_o V}} \sum_{\vec{k}} \tilde{e}_k \left[\hat{q}_k \cos\left(\vec{k}\cdot\vec{r} - \omega_k t\right) - \frac{\hat{p}_k}{\omega_k} \sin\left(\vec{k}\cdot\vec{r} - \omega_k t\right) \right] \tag{3.6.3}$$

$$\vec{E}(\vec{r},t) = \frac{-1}{\sqrt{\varepsilon_o V}} \sum_{\vec{k}} \tilde{e}_k \, \omega_k \left[\hat{q}_k \sin\left(\vec{k}\cdot\vec{r} - \omega_k t\right) + \frac{\hat{p}_k}{\omega_k} \cos\left(\vec{k}\cdot\vec{r} - \omega_k t\right) \right] \tag{3.6.4}$$

$$\vec{B}(\vec{r},t) = \frac{-1}{\sqrt{\varepsilon_o V}} \sum_{\vec{k}} \vec{k} \times \tilde{e}_k \left[\hat{q}_k \sin\left(\vec{k}\cdot\vec{r} - \omega_k t\right) + \frac{\hat{p}_k}{\omega_k} \cos\left(\vec{k}\cdot\vec{r} - \omega_k t\right) \right] \tag{3.6.5}$$

The position and momentum operators are related to the quadrature components of the fields (need to double check the minus signs). The position and momentum operators are not related to the polarization of the electromagnetic wave. Similarly, neither operator can be identified solely with just one of the fields. At $\vec{r} = 0$ and t=0 the electric field is directly proportional to the momentum operator $\hat{p}$ while at a later time the electric field is directly proportional to the position operator $\hat{q}$. Similarly, changing the point of observation $\vec{r}$ also changes the relation between the electric field and the operators.

Section 3.7: Rotation Operator for the Quantized Electric Field

Classical travelling-wave electric fields have an arbitrary origin of time which is equivalent to saying the initial phase ϕ of the field can be set to an arbitrary value. There are occasions when we would like an operator that "rotates" the electric field operator to an arbitrary phase such as

$$\hat{E} = i\sum_k \sqrt{\frac{\hbar\omega_k}{2\varepsilon_o V}}\left[\hat{b}_k e^{i\vec{k}\cdot\vec{r}-i\omega_k t-i\theta_k} - \hat{b}_k^+ e^{-i\vec{k}\cdot\vec{r}+i\omega_k t+i\theta_k}\right]\tilde{e}_k$$

where each mode can be separately rotated. Besides being interesting and important in its own right, the single-mode rotation operator

$$\hat{R}_k(\theta) = e^{-i\theta_k \hat{N}_k} = e^{-i\theta_k \hat{b}_k^+\hat{b}_k}$$

allows an interchange quadrature components and facilitates the discussion of the Wigner probability function. The rotation of all modes can be simultaneously implemented by applying the rotation operator

$$R = \sum_k R_k$$

For now, concentrate on rotating the single mode

$$\hat{E} = i\sqrt{\frac{\hbar\omega}{2\varepsilon_o V}}\left[\hat{b}e^{i\vec{k}\cdot\vec{r}-i\omega t} - \hat{b}^+ e^{-i\vec{k}\cdot\vec{r}+i\omega t}\right]$$

The rotation operator is defined to be

$$\hat{R}(\theta) = e^{-i\theta\hat{N}} = e^{-i\theta\hat{b}^+\hat{b}}$$

which is obviously unitary with $\hat{R}^+(\theta) = \hat{R}^{-1}(\theta) = \hat{R}(-\theta)$. The phase parameter θ is real. The number operator, which is the conjugate variable to the phase operator, is defined as usual by $\hat{N} = \hat{b}^+\hat{b}$. The number operator is the generator of phase rotations. The rotated field is defined by a similarity transformation

$$\hat{E}_R = \hat{R}^+\hat{E}\hat{R}$$

Either the state can be rotated $\hat{R}^+|\psi\rangle$ or the operator can be rotated using the similarity transformation (but we should not do both). To apply the similarity transformation, it is necessary to know how the rotation affects the creation and annihilation operators.

We first show that

$$\hat{R}^+\hat{b}\hat{R} = e^{i\theta\hat{b}^+\hat{b}}\hat{b}e^{-i\theta\hat{b}^+\hat{b}} = \hat{b}e^{-i\theta} \quad \text{and} \quad \hat{R}^+\hat{b}^+\hat{R} = \hat{b}^+e^{i\theta} \tag{3.7.1}$$

The operator expansion theorem

$$e^{x\hat{A}}\hat{B}e^{-x\hat{A}} = \hat{B} + \frac{x}{1!}[\hat{A},\hat{B}] + \frac{x^2}{2!}\left[\hat{A},[\hat{A},\hat{B}]\right] + \ldots$$

with $x = i\theta$, $\hat{A} = \hat{b}^+\hat{b}$ and $\hat{B} = \hat{b}$ provides

$$\begin{aligned} e^{i\theta\hat{b}^+\hat{b}}\hat{b}e^{-i\theta\hat{b}^+\hat{b}} &= \hat{b} + \frac{i\theta}{1!}[\hat{b}^+\hat{b},\hat{b}] + \frac{(i\theta)^2}{2!}\left[\hat{b}^+\hat{b},[\hat{b}^+\hat{b},\hat{b}]\right] + \ldots \\ &= \hat{b} - i\theta\hat{b} + \ldots \\ &= \hat{b}e^{-i\theta} \end{aligned}$$

where the results used $\left[\hat{b},\hat{b}^{+}\right]=1$ and $\left[\hat{A}\hat{B},\hat{C}\right]=\hat{A}\left[\hat{B},\hat{C}\right]+\left[\hat{A},\hat{C}\right]\hat{B}$. Also notice that the previous two relations are adjoints of each other.

Now it's easy to see that the single-mode rotation operator rotates the phase of the electric field.

$$\begin{aligned}\hat{E}_R &= \hat{R}^{+}\hat{E}\hat{R} = i\sqrt{\frac{\hbar\omega}{2\varepsilon_o V}}\left[\left(\hat{R}^{+}\hat{b}\hat{R}\right)e^{i\vec{k}\cdot\vec{r}-i\omega t}-\left(\hat{R}^{+}\hat{b}^{+}\hat{R}\right)e^{-i\vec{k}\cdot\vec{r}+i\omega t}\right]\\ &= i\sqrt{\frac{\hbar\omega}{2\varepsilon_o V}}\left[\left(\hat{b}e^{-i\theta}\right)e^{i\vec{k}\cdot\vec{r}-i\omega t}-\left(\hat{b}^{+}e^{i\theta}\right)e^{-i\vec{k}\cdot\vec{r}+i\omega t}\right] \qquad (3.7.2)\\ &= i\sqrt{\frac{\hbar\omega}{2\varepsilon_o V}}\left[\hat{b}e^{i\vec{k}\cdot\vec{r}-i\omega t-i\theta}-\hat{b}^{+}e^{-i\vec{k}\cdot\vec{r}+i\omega t+i\theta}\right]\end{aligned}$$

as required.

Next, the rotation operator can be used to interchange the quadrature terms in the single-mode electric field operator

$$\hat{E}=\frac{-\omega}{\sqrt{\varepsilon_o V}}\left[\hat{q}\sin\left(\vec{k}\cdot r-\omega t\right)+\frac{\hat{p}}{\omega}\cos\left(\vec{k}\cdot r-\omega t\right)\right]$$

The procedure requires that the rotation operator be applied to the genralized "position" and "momentum" operators. Equation 3.7.2 shows that the replacement $\vec{k}\cdot\vec{r}-\omega t\rightarrow\vec{k}\cdot\vec{r}-\omega t-\theta$ can be made. Setting $\theta=\frac{\pi}{2}$ and using the relations $\cos\left(\phi-\frac{\pi}{2}\right)=\sin(\phi)$ and $\sin\left(\phi-\frac{\pi}{2}\right)=\cos(\phi)$ provides

$$\hat{E}=\frac{-\omega}{\sqrt{\varepsilon_o V}}\left[\hat{q}\cos\left(\vec{k}\cdot r-\omega t\right)+\frac{\hat{p}}{\omega}\sin\left(\vec{k}\cdot r-\omega t\right)\right]$$

Now the position and momentum operators correspond to a horizontal and vertical axis respectively.

Chapter 4: Quantum States

The previous chapter discusses the quantization of the electromagnetic fields and Hamiltonian. Operators replace the classical dynamical variables. The structure of the theory requires that the field operators and Hamiltonian include a sum over all possible modes of the system. The number of potential modes is determined by boundary conditions. For example, for a normalization volume the size of the known universe, there are a very large number of possible modes in which case, the summation over the modes can be replaced with an integral. The last chapter also shows that the quantized fields have both the wave and particle aspects built into them. Creation and annihilation operators represent the particle aspect.

The structure of quantum theory requires that operators act on wave functions in a Hilbert space. The operators can, in fact, be decomposed in terms of the basis vectors for the Hilbert space if desired. The wave functions (or states) contain the specific information on the system. For example, there might only be a finite number of photons to populate an infinite number of possible states. The act of observing the fields and energy is equivalent to operating with the quantized fields on the state vectors as discussed in Chapter 2. The average value obtained by repeated observation of a physical quantity is mathematically obtained by forming the expectation value of the corresponding operator.

Fock states are basis vectors in a Hilbert space which are eigenvectors of the number operator (or, equivalently, the Hamiltonian). These states have an exact number of photons in each modes but the phase of the photons is unknown. Although the energy stored in a Fock state is known, the expected value of the electric field is zero (maybe contrary to intuition). The average is zero because the phase of the field is randomly distributed over all possible phase angles $\phi(t)$ in $e^{ikz-i\omega t+\phi(t)}$. In short, the Fock state is not an eigenstate of the fields. A problem with measuring the field can be anticipated because, for each mode, there are two quadrature terms that do not commute and hence lead to an uncertainty relation. Perhaps the energy is the more correct quantity to measure because it is more physical than the fields. Recall that the fields were originally defined, in the 1700s-1800s, in terms of the work done on a positive test charge in moving it from one location to another. The situation can be clarified somewhat by forming linear combinations of the Fock states.

The annihilation operator has the coherent state as its eigenstate. The coherent state is a linear combination of all Fock states. The coherent state describes the *average* amplitude and *average* phase of an electromagnetic wave (as opposed to the exact number of photons in each mode). *Repeated* measurement of the electric field on the coherent state gives a sinusoidal wave, which has the average amplitude and phase characterizing that coherent state (i.e., the electric field can be pictured as sines, cosines or travelling waves similar to classical optics). However, because of the non-commuting quadrature terms in the electric field, the amplitude and phase cannot be simultaneously and precisely measured. Therefore, the amplitude and phase take on a range of values which end-up being described by Poisson statistics (shot noise). Moving charges and lasers generate coherent states.

The squeezed state is built upon the coherent state. Fluctuations are removed from one parameter (such as the number of photons) but reappear in the conjugate

parameter (such as phase) or vice versa. A perfectly number-squeezed state (i.e., all noise removed from the amplitude) is a Fock state. The squeezed state can be generated by a number of devices and methods including semiconductor lasers, masers, parametric amplifiers and four-wave mixing.

Section 4.1: Introduction to Fock States

This section discusses the more general notion of Fock states and shows how they originate from single-particle basis-state wavefunctions. Fock states are intuitive for counting photons and there are many accessible properties. Fock states have a definite number of particles but the phases of those particles are completely unknown. This section shows that the average electric field is zero when the system is described by a Fock state. Later sections in this chapter show that laser light is best described by the coherent state which describes the amplitude of the electric field.

Topic 4.1.1: Concepts for Creation/Annihilation Operators and Fock States

The creation and annihilation operators are similar to the harmonic oscillator raising and lowering operators discussed in Chapter 2. For the harmonic oscillator, the ket $|n\rangle = |u_n\rangle$ represents a particular energy eigenstate for a particle. However, each ket $|n\rangle$ can be *re-interpreted* as the number of available quanta in the oscillator; i.e., there is a 1-1 correspondence between the energy eigenstate occupied by a particle and the number of available quanta in that state. The ladder operators map an energy eigenstate into the next state in the sequence of eigenstates; this mapping is equivalent to adding or subtracting a quantum of energy. Mathematically speaking, the ladder operators act on a vector space with basis vectors $\{|n\rangle = u_n\}$ and map one basis vector into another $\phi_n \leftrightarrow \phi_{n+1}$ according to

$$a^+|n\rangle = \sqrt{n+1}\,|n+1\rangle \qquad a|n\rangle = \sqrt{n}\,|n-1\rangle$$

or equivalently,

$$a^+ = \sum_{n=0}^{\infty} \sqrt{n+1}\,|n+1\rangle\langle n| \qquad a = \sum_{j=1}^{\infty} \sqrt{n}\,|n-1\rangle\langle n|$$

Creation "b^+" and annihilation "b" operators are somewhat different (the symbol "b" is used as a reminder that photons are bosons) than the ladder operators. The creation and annihilation operators act on Fock space. Fock states are (in a sense) a mathematical device for tracking (i.e., "book-keeping") the number of particles in a given mode of the system. The number of particles is considered to be exact (i.e., the standard deviation of the number operator evaluated in a Fock state is zero). A creation operator places a particle in one (out of many) modes by creating the particle from the vacuum state. An annihilation operator removes a particle from a mode. The particles can be either Fermions (particles with half integer spin such as electrons) or they can be bosons (particles with integer spin such as photons). Figure 4.1.1 shows a ket in Fock space as consisting of "buckets" which hold the particles.

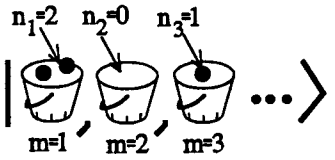


Figure 4.1.1: Fock states as "book-keeping" for the number of photons in each mode.

As an example, consider *the Fabry-Perot cavity* shown in the Figure 4.1.2 with the m=1 and m=2 optical modes (the sine waves represent the electric field). Of course, there are more than two optical modes but they not drawn. The mirrors (drawn as black

boxes) provide "boundary conditions" and give rise to a discrete spectrum for the wavelength λ_m which characterizes the allowed modes (eigenfunctions).

$$\lambda_m = 2L, L, \frac{2L}{3} \ldots \frac{2L}{m} \qquad m = 1,2,3\ldots$$

Note that the mode number "m" is nonzero for this example. The energy of a *single* photon is given by

$$E_m = \frac{hc}{\lambda_m} = \left(\frac{hc}{2L}\right) m$$

where Planck's constant is $h = 6.63 \cdot 10^{-34}$ and the speed of light in vacuum is $c = 3 \cdot 10^{8}$ in MKS units. The important point here is that the eigenfunctions

$$u_m(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{m\pi}{L} x\right)$$

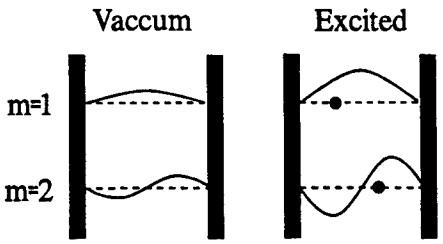


Figure 4.1.2: Classical electric field vs. distance for photons in the Fabry-Perot cavity. For a Fock state, the figure is inaccurate in that it shows a well-defined phase contrary the Heisenberg uncertainty relation.

corresponding to the energy E_m represent the modes of the Fabry-Perot cavity. Notice that the integer "m" is never 0.

Now to change points of view. Consider a direct product space denoted by

$$|n_1, n_2, n_3 \ldots\rangle$$

Suppose that the first position in the ket $|\quad\rangle$ stands for a mode with wave vector $\vec{k}_1$ (i.e., wavelength $\lambda = \lambda_1$). The symbol n_1 gives the number of photons in mode number 1. Similarly n_2 is the number of photons in the mode with wave vector $\vec{k}_2$ with wavelength $\lambda_2 = \frac{2L}{n} = L$. Figure 4.1.2 shows that n_1=n_2=1. The state vector is $|1,1,000\ldots\rangle$. The total *accessible* energy stored in the cavity is

$$E_m = \frac{hc}{\lambda_m} = \left(\frac{hc}{2L}\right) m$$

to give

$$E_{tot} = E_1 + E_2 = \frac{hc}{2L} + 1\left(\frac{hc}{2L} 2\right) = \frac{3}{2}\frac{hc}{L}$$

State vectors $|n_1, n_2, \ldots\rangle$ that explicitly keep track of the number of particles in a mode are called Fock states. The position in the Fock vector corresponds to a given mode which can include polarization and wavelength for photons. As an important note, the vacuum energy is not included in E_{tot} above, which is where the factor of ½ becomes important for the photon Hamiltonian in Equation 3.4.8.

If the system has no photons, then the system is in the so-called vacuum state which is written as

$$|0,0,0,\ldots\rangle = |0\rangle$$

If there is only one allowed *mode*, say the mode characterized by the wavelength λ_1, then the Fock vectors would have only one position

$$|n_1\rangle = n_1 \text{ particles in the mode } \lambda_1$$

$$|0\rangle = \text{no particles in the mode m=1}$$

This looks very similar to the notation used in conjunction with the ladder operators.

Now define creation operators $\hat{b}_i^+ = \hat{b}_i^+(0)$ and annihilation operators $\hat{b}_i = \hat{b}_i(0)$ as

$$b_i^+ |n_1, n_2, \ldots, n_i, \ldots\rangle = \sqrt{n_i + 1}\, |n_1, n_2, \ldots, n_i + 1, \ldots\rangle$$

$$b_i |n_1, n_2, \ldots, n_i, \ldots\rangle = \sqrt{n_i}\, |n_1, n_2, \ldots, n_i - 1, \ldots\rangle$$

where b_i^+ creates a particle in mode "i" and b_i removes a particle.

If the initial state is the quantum mechanical vacuum then b_i^+ creates a particle mode #i according to

$$b_i^+ \left|0,0,\ldots,\underset{\#i}{0},\ldots\right\rangle = \left|0,0,\ldots,\underset{\#i}{1},\ldots\right\rangle$$

The annihilation operator can remove a particle from mode #i according

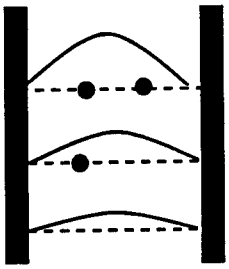
$$b_i \left|0,0,\ldots,\underset{\#i}{1},\ldots\right\rangle = \left|0,0,\ldots,\underset{\#i}{0},\ldots\right\rangle$$

To move a particle from mode #1 to mode #2, for example, it is first necessary to destroy the particle in mode 1 and create it in mode 2 as follows

$$b_2^+ b_1 |1,0,0,\ldots\rangle = |0,1,0,0\ldots\rangle$$

Notice that for the harmonic oscillator, a single ladder operator a^+ can move a particle from mode $|1\rangle$ to $|2\rangle$ according to $a^+|1\rangle = \sqrt{2}|2\rangle$. But Fock states require two operators so that $a^+ \sim b_2^+ b_1$. Of course, something like this should be expected since $a^+ = \sum \sqrt{n+1}|n+1\rangle\langle n|$ and the bra $\langle n|$ acts similar to the annihilation operator b_n while $|n+1\rangle$ is roughly equivalent to the creation operator b_{n+1}^+.

The meaning of adding a photon to the Fabry-Perot cavity mode can be give a very rough physical picture of the cavity. Adding a photon is equivalent to adding energy. The amplitude of the electric field is directly related to the energy in an electromagnetic wave. Therefore adding a photon must increase the amplitude as roughly shown in the Figure 4.1.3 for a single mode. For a Fock state, the number of photons in a mode is known exactly; however, the phase of the wave is completely unknown. The phase of the wave for the Fabry-Perot cavity does not refer to the shape of the sine waves versus position between the mirrors. Rather the phase ϕ in $\sin(kx)e^{i\omega t+\phi}$ (or equivalently, the origin of time) refers to whether the peak of the wave is at the top (as drawn), at the bottom (180° phase shift), or somewhere in between. There is a Heisenberg uncertainty relation between the particle number "n" and "ϕ" of the form



Single Mode

Figure 4.1.3: Plot of the electric field for a single optical mode with either 0, or 1 or 2 photons. Note the inaccuracy of the figure: the phase should not be well defined.

$$\Delta n\, \Delta\phi \geq \frac{1}{2}$$

where recall that Δ represents the standard deviation. Actually this relation is more correctly given in terms of $\sin(\phi)$ rather than ϕ.

Notice in Figures 4.1.2 and 4.1.3 that the even though a mode might not have any photons, there is still an electric field! The electric field is related to the zero point motion similar to that for the harmonic oscillator in Chapter 2. Even though there isn't any available energy, there is still a fluctuating electric field. These quantum fluctuations lead to spontaneous emission. The Casimir effect is one example of a real effect due to vacuum modes where the vacuum fluctuations can be experimentally observed to move two metal plates toward each other.

Topic 4.1.2: Including the Photon Polarization

So far, the polarization of the photon has not been included in the Fock state description. The Coulomb gauge gives rise to traveling waves with only two possible polarization directions which are perpendicular to the direction of propagation. Assume the two basic polarization directions $\tilde{e}_{ks}$ (S=1,2) for each wavevector $\vec{k}$. Explcitly, if the wave propagates along the z-direction, then the S=1 polarization mode is along $\hat{x}$ and the other S=2 mode is along $\hat{y}$. Circular polarization unit vectors can also be used rather than the plane-wave polarization vectors.

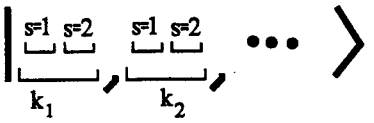


Figure 4.1.4: A Fock state includes the polarization.

The creation and annihilation operators create and remove photons from a mode characterized by a given wave vector and given polarization. These operators are subscripted by both $\vec{k}$ and s.

$$\hat{b}_{is}^{+}\left|n_{1s},n_{2s},\ldots,n_{is},\ldots\right\rangle = \sqrt{n_{is}+1}\left|n_{1s},n_{2s},\ldots,n_{is}+1,\ldots\right\rangle$$

$$\hat{b}_{is}\left|n_{1s},n_{2s},\ldots,n_{is},\ldots\right\rangle = \sqrt{n_{is}}\left|n_{1s},n_{2s},\ldots,n_{is}-1,\ldots\right\rangle$$

Usually, the polarization index (sometimes called the spin index) is suppressed.

The next couple of topics point out the use of Fock space and related operators for Bosons and Fermions. Particles with half-integer spin (such as electrons) constitute the set of Fermions while particles with integer spin (such as photons or phonons) make up the class of Bosons.

Topic 4.1.3: Mathematics of Boson Fock States and Creation/Annihilation Operators

Bosons are particles with integer spin (0,1,2,...), such as photons and phonons. Any number of them can occupy a given mode. For Fermions, the commutation relations for the creation and annihilation operators control the number of possible particles per state. For a given set of modes, each Fock state (i.e., a given n_1, n_2, n_3 ...) is a basis vector. The *complete* set of basis vectors is given by

$$\left\{\left|n_1,n_2,n_3,\ldots\right\rangle\right\}$$

where each n_i can range from zero up to an *infinite* number of particles in the system. The reader should keep in mind that the positions in the ket refer to the different possible modes in the system and therefore, the *number of positions* is equal to the total number of available modes. The orthonormality relation is

$$\langle n_1, n_2, \ldots | m_1, m_2, \ldots \rangle = \delta_{n_1 m_1} \delta_{n_2 m_2} \cdots$$

and the closure relation is

$$\sum_{n_1, n_2 \ldots = 0}^{\infty} |n_1, n_2 \ldots\rangle \langle n_1, n_2 \ldots| = \hat{1}$$

The creation $\hat{b}_k^+$ and annihilation $\hat{b}_k$ operators increase or decrease the number of particles in a particular mode according to

$$\hat{b}_k^+ |n_1, n_2, \ldots, n_k, \ldots\rangle = \sqrt{n_k + 1}\, |n_1, n_2, \ldots, n_k + 1, \ldots\rangle$$

$$\hat{b}_k |n_1, n_2, \ldots, n_k, \ldots\rangle = \sqrt{n_k}\, |n_1, n_2, \ldots, n_k - 1, \ldots\rangle$$

with commutation relations given by

$$\left[\hat{b}_k, \hat{b}_K\right] = 0 = \left[\hat{b}_k^+, \hat{b}_K^+\right] \quad \text{and} \quad \left[\hat{b}_k, \hat{b}_K^+\right] = \delta_{kK}$$

where the subscripts must also include the spin (i.e., polarization).

The mode-number operator

$$\hat{N}_k = \hat{b}_k^+ \hat{b}_k$$

gives the number of paricles in the state "k". The Fock states are eigenstates of the number operator

$$\hat{N}_k |n_1, \ldots, n_k, \ldots\rangle = n_k |n_1, \ldots, n_k, \ldots\rangle$$

The total number of particles in a Fock state can be found by using the total-number operator

$$\hat{N} = \sum_i \hat{N}_i$$

so that

$$\hat{N} |n_1, \ldots, n_k, \ldots\rangle = \left(\sum_i \hat{N}_i\right) |n_1, \ldots, n_k, \ldots\rangle = \left(\sum_i n_i\right) |n_1, \ldots, n_k, \ldots\rangle$$

The number operators have a "sharp" value for the Fock states which means their standard deviation is zero. The standard deviation is zero for any operator $\hat{O}$ evaluated in its eigenstate $|\phi\rangle$ (i.e., $\hat{O}|\phi\rangle = \phi |\phi\rangle$) as can be seen by starting with

$$\overline{O} = \langle \phi | \hat{O} | \phi \rangle = \phi \langle \phi | \phi \rangle = \phi$$

so that

$$\sigma_{\hat{O}}^2 = \langle \phi | \left(\hat{O} - \overline{O}\right)^2 | \phi \rangle = \langle \phi | \hat{O}^2 | \phi \rangle - \overline{O}^2 = \phi^2 \langle \phi | \phi \rangle - \phi^2 = 0$$

Physically "sharp values" means that repeated measurements produce only one value (i.e., the measurement does not interfere with the system).

Topic 4.1.4: Mathematics of Fermion Fock States and Creation/Annihilation Operators

Fermions are particles, such as electrons, that have half-integral spin

$$\frac{1}{2},\frac{3}{2},\frac{5}{2}\cdots$$

The commutation relations for Fermions demonstrate the Pauli exclusion principle which mandates that only a single Fermion can occupy a single state at one time. The Fermion creation $\hat{f}_k^+$ and annihilation $\hat{f}_k$ operators obey *anticommutation* relations given by

$$\left[\hat{f}_k,\hat{f}_K\right]_+ = 0 = \left[\hat{f}_k^+,\hat{f}_K^+\right]_+ \quad \text{and} \quad \left[\hat{f}_k,\hat{f}_K^+\right]_+ = \delta_{kK}$$

where the anticommutator is defined by

$$\left[\hat{A},\hat{B}\right]_+ = \hat{A}\hat{B}+\hat{B}\hat{A}$$

Notice the anticommutator uses a "+" sign which makes all the difference for the particle statistics.

Let's try to create two Fermions in a single state (neglecting all but one mode).

$$\hat{b}^+\hat{b}^+|0\rangle$$

The anticommutation relation for the creation operator provides

$$0=\left[\hat{b}^+,\hat{b}^+\right]=\hat{b}^+\hat{b}^+ +\hat{b}^+\hat{b}^+ = 2\hat{b}^+\hat{b}^+$$

so that the two particle Fermion ket becomes

$$\hat{b}^+\hat{b}^+|0\rangle = 0$$

The anticommutation relations for Fermions therefore lead to the Pauli exclusion principle.

Section 4.2: Fock States as Solutions to the EM Schrodinger Equation

Schrodinger's equation can be written in terms of either (1) the generalized coordinates or (2) the Fock states. As the previous chapter discusses, the two forms of the Hamiltonian are

$$\hat{H} = \sum_k \left[\frac{p_k^2}{2} + \frac{\omega^2 q_k^2}{2} \right] \qquad \hat{H} = \sum_k \hbar\omega_k \left(\hat{N}_k + \frac{1}{2} \right) \tag{4.2.1}$$

For the first case, the Schrodinger equation is stated in terms of the coordinate representation of the momentum operator. This version of the Schrodinger equation has wave function solutions that depends on the generalized coordinates. The wave functions gives the probability amplitude that a measurement will give a particular electric field amplitude. The second form of the Schrodinger equation is the most commonly employed and it is the simplest to solve and manipulate (in fact, it's already solved in the previous section!). The appendices show how the multi-particle (many body) wavefunction leads to the Fock states.

Topic 4.2.1: Coordinate Representation of Boson Wavefunctions

In this topic, the solution to the EM Schrodinger equation obtained by separation of variables is stated. The creation and annihilation operators are used to find the actual eigenfunctions.

In the coordinate representation, the generalized momentum operator is replaced with

$$\hat{p}_k \rightarrow \frac{\hbar}{i} \frac{\partial}{\partial q_k}$$

so that the Hamiltonian becomes

$$\hat{H} = \sum_k \left[\frac{p_k^2}{2} + \frac{\omega^2 q_k^2}{2} \right] \rightarrow \hat{H} = \sum_k \left[-\frac{\hbar^2}{2} \frac{\partial^2}{\partial q_k^2} + \frac{\omega^2 q_k^2}{2} \right]$$

The wavefunction must depend on the independent coordinates q_k so that Schrodinger's equation for light can be written as

$$\sum_{k=1}^{N} \left[-\frac{\hbar^2}{2} \frac{\partial^2}{\partial q_k^2} + \frac{\omega^2 q_k^2}{2} \right] \Psi(q_1, q_2, \ldots, q_N, t) = i\hbar \frac{\partial}{\partial t} \Psi(q_1, q_2, \ldots, q_N, t) \tag{4.2.2}$$

where N is the number of photons. The polarization and wave vector direction are ignored. Separate variables provides

$$\psi(q_1, q_2, \ldots, q_N, t) = u_{E_1}(q_1)\, u_{E_2}(q_2) \ldots u_{E_N}(q_N)\, T(t)$$

and each basis function $u_{E_k}(q_k)$ satisfies a time-independent Shrodinger's equation with a form similar to the harmonic oscillator

$$\left(-\frac{\hbar^2}{2} \frac{\partial^2}{\partial q_k^2} + \frac{\omega^2 q_k^2}{2} \right) u_{E_k}(q_k) = E_k\, u_{E_k}(q_k) \tag{4.2.3}$$

Before proceeding with the solution, it is useful develop an intuitive understanding of a wavefunction such as $u_{E_k}(q_k)$. On the one hand, Equation 4.2.3 does *not* reference the spatial position of the photon; it suggests that the focus should be on the amplitude of the oscillations of the electromagnetic field. Figure 4.2.1 shows that the harmonic motion of the wave is associated with the oscillation of the field about its "equilibrium." This should be compared with the harmonic oscillator wave functions in Chapter 2, such as $\psi(x)$, which gives the probability amplitude of finding a massive particle at the *spatial postion* "x". As previously mentioned, the coordinate q_k is similar to an electric field for the k^{th} mode. Intuitively, the wave function $u(q_k)$ gives the probability amplitude that a photon in mode "k" will be found with an electric field amplitude of q_k. This probability amplitude interpretation for $u(q_k)$ can be seen from, for example, Equation 3.6.2 in Section 3.6

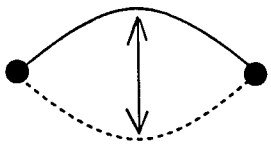


Figure 4.2.1: The arrow indicates the harmonic oscillation in the field.

$$\vec{E}(\vec{r},t)=\frac{-1}{\sqrt{\varepsilon_o V}}\sum_{\vec{k}}\tilde{e}_k\,\omega_k\left[\hat{q}_k\sin\left(\vec{k}\cdot\vec{r}-\omega_k t\right)+\frac{\hat{p}_k}{\omega_k}\cos\left(\vec{k}\cdot\vec{r}-\omega_k t\right)\right]$$

by considering a point in time and space such that "kz-ωt=π/2". This provides

$$\vec{E}\sim\frac{-1}{\sqrt{\varepsilon_o V}}\sum_{\vec{k}}\tilde{e}_k\,\omega_k\hat{q}_k$$

Therefore, the wave function $u_{E_k}(q_k)$ is the probability amplitude of finding a particular electric field amplitude as represented by "q_k".

Returning to the solution of Equation 4.2.3, the Hamiltonian is similar to that for the harmonic oscillator. Chapter 2 shows that the eigenvectors consist of exponentials and Hermite polynomials. The eigenvalues are

$$E_k=\hbar\omega_k\left(n_k+\frac{1}{2}\right)$$

Therefore, the total energy for all N particles is given by

$$E=\sum_{k=1}^{N}\hbar\omega_k\left(n_k+\frac{1}{2}\right)$$

The eigenfunctions with a time dependent phase factor is given by

$$\psi(q_1,q_2,\ldots,q_N,t)=u_{E_1}(q_1)\,u_{E_2}(q_2)\ldots u_{E_N}(q_N)\,T(t)=u_{E_1}(q_1)\,u_{E_2}(q_2)\ldots u_{E_N}(q_N)\,e^{-itE/\hbar}$$

A general wave function in the multidimensional Hilbert space can be written as

$$\psi(q_1,q_2,\ldots,q_N,t)=\sum_{E_1,E_2\ldots E_N}\beta(E_1,E_2,\ldots,E_N,t)\,u_{E_1}(q_1)\,u_{E_2}(q_2)\ldots u_{E_N}(q_N)$$

where β includes the phase factor. Appendix 4 shows that this equation can actually be rearranged to provide the Fock states. Here, a solution is sought only in the simple case of a single mode by using creation and annihilation operators.

We want to find the *single mode* wave function $\psi_n(q)$ *where "n" stands for the number of photons in the mode* and the wavefunction satisfies Schrodinger's equation for light; that is

$$\langle q|n\rangle=u_n(q)$$

and

$$\hat{H}\,u_n(q) = E_n u_n(q)$$

Similar to the Chapter 2 harmonic oscillator, the partial differential equation can be solved using a series method. However, it is simpler to find the wave functions $u_n(q)$ by repeatedly applying the annihilation and creation operators to the vacuum state. First apply the destruction operator to the vacuum state

$$\hat{b}|0\rangle = 0$$

or using the position and momentum representation given in Equation 3.4.10

$$\left(\frac{\omega\ \hat{q}}{\sqrt{2\hbar\omega}} + \frac{i\,\hat{p}}{\sqrt{2\hbar\omega}}\right)|0\rangle = 0$$

where ω is the angular frequency for light in the mode. Operating with the coordinate space operator $\langle q|$ (i.e., projecting into coordinate space) provides

$$\langle q|\left(\frac{\omega\ \hat{q}}{\sqrt{2\hbar\omega}} + \frac{i\,\hat{p}}{\sqrt{2\hbar\omega}}\right)|0\rangle = 0$$

or, inserting the coordinate representation of the operators, provides

$$\left(\frac{\omega\ q}{\sqrt{2\hbar\omega}} + \frac{\hbar}{\sqrt{2\hbar\omega}}\frac{\partial}{\partial q}\right)\langle q|0\rangle = 0$$

Therefore, a simple first order differential equation for $\langle q|0\rangle = u_0(q)$ must be solved. The solution is

$$u_0(q) = \left(\frac{\omega}{\pi\hbar}\right)^{1/4} \exp\left[-\frac{\omega q^2}{2\hbar}\right] \qquad (4.2.4)$$

where the constant comes from the normalization condition. Equation 4.2.4 gives the probability amplitude of finding the electric field represented by "q" in the vacuum state (i.e., zero photons).

Just like the harmonic oscillator in Chapter 2, it is possible to find all of the ensuing wavefunctions by repeatedly applying the creation operator. For the first excitation of the mode (i.e., n=1 corresponding to a single photon)

$$|1\rangle = \hat{b}^+|0\rangle$$

Operating with the coordinate space projector and substituting the coordinate representation of the creation operator provides

$$u_1(q) = \left(\frac{\omega\ q}{\sqrt{2\hbar\omega}} - \frac{\hbar}{\sqrt{2\hbar\omega}}\frac{\partial}{\partial q}\right)u_0(q)$$

A result similar to the massive-particle harmonic oscillator is found

$$u_1(q) = \sqrt{\frac{2\omega}{\hbar}}\left(\frac{\omega}{\pi\hbar}\right)^{1/4} q\exp\left(-\frac{\omega q^2}{2\hbar}\right) \qquad (4.2.5)$$

The Hermite polynomials start to emerge with increasing photon number n.

Figure 4.2.2 shows an example

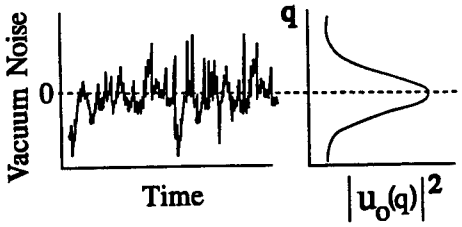


Figure 4.2.2: Comparing the coordinate "q" with the measured noise in the electric field for the vacuum.

measurements of the "electric field amplitude" q for the vacuum state $|0\rangle$. The probability density is plotted (sideways) next to the measured signal. Recall that the probability density is the modulus squared of the probability amplitude represented by the wave function. The figure shows the greatest excursions from the average occurs only a few times; the probability is smallest for these "positions" q. For a fixed number of photons (such as n=0) the "electric field q" can be observed with a variety of amplitdues (similar comments apply to the "p-field"). Therefore, the "electric field amplitude q" is not fixed even though the number of photons is fixed. Although easier to find using the number representation of the Fock states, the average "position q" (i.e., electric field) is found by integration to be

$$\langle u_0(q)|q|u_0(q)\rangle = \int_{-\infty}^{\infty} dq\, u_0^*(q)\, q\, u_0(q) = 0$$

The standard deviation of q, namely $\sigma_q = \Delta q$ can be found by calculating

$$\sigma_q{}^2 = \langle u_0(q)|\,(q-\underline{q})^2\,|u_0(q)\rangle = \langle u_0(q)|\,q^2\,|u_0(q)\rangle$$

The average "momentum p" can be calculated

$$\langle u_0(q)|\,p\,|u_0(q)\rangle = \langle u_0(q)|\,\frac{\hbar}{i}\frac{\partial}{\partial q}\,|u_0(q)\rangle = \int_{-\infty}^{\infty} dq\, u_0^*(q)\,\frac{\hbar}{i}\frac{\partial}{\partial q}\,u_0(q) = 0$$

and similarly the standard deviation of p, namely $\sigma_p = \Delta p$ can be found by calculating

$$\sigma_p{}^2 = \langle u_0(q)|\left(\frac{\hbar}{i}\frac{\partial}{\partial q}\right)^2|u_0(q)\rangle$$

If the calculations are carried through (refer to the next sections), the Heisenberg uncertainty relation is found

$$\Delta q\,\Delta p = \frac{\hbar}{2} \quad \text{for the vacuum state}$$

The figure at right compares the vacuum state with the n=1 Fock state which has one photon. The average electric field is zero for each $\langle q\rangle = 0$ but the standard deviation for the first excited state is larger than for the vacuum. Multiple measurements of the electric field for the first excited state does not yield any values with q=0; however, non-zero fields are found for "q" near the peaks of the wavefunction. For all Fock states (multiple photons) a Heisenberg uncertainty relation obtains

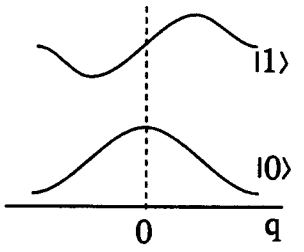


Figure 4.2.3: Comparing the photon wave function for a single mode with either 0 or 1 photon.

$$\Delta q\,\Delta p \geq \frac{\hbar}{2}$$

with the "equal sign" applying only to the vacuum. It is clear from the expression for the quantized field

$$\vec{E}(\vec{r},t) = \frac{-1}{\sqrt{\varepsilon_o V}}\sum_{\vec{k}} \tilde{e}_k\,\omega_k\left[\hat{q}_k\sin(\vec{k}\cdot\vec{r}-\omega_k t) + \frac{\hat{p}_k}{\omega_k}\cos(\vec{k}\cdot\vec{r}-\omega_k t)\right]$$

that the two quadrature terms do not commute $[\hat{q}_k,\hat{p}_k] = i\hbar$, which results in the corresponding Heisenberg uncertainty relation. It is not possible to simultaneously

measure both components of the electric field with infinite accuracy. Every time a measurement is made of the field, the quadrature components have a different value. This is where the Wigner function becomes important because it provides a quasi-classical probability density for the two coordinates q,p.

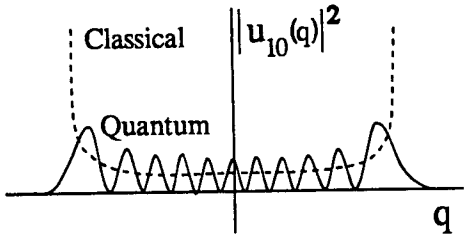


Figure 4.2.4: Comparing the quantum mechanical probability for a particular field amplitude (10 photons in the mode) with the classical counterpart.

As a note, the figure at right compares the "classical probability" for finding an electric field with a given value versus that for the quantized field (n=10 photons in a single mode). The larger is the number of photons in a mode, the closer the quantum electric field is to the classical field. That is, the noise in the field is not as important and the S/N increases.

Topic 4.2.2: Fock States as Energy Eigenstates

This topic solves Schrodinger's equation in the Fock number representation. As a summary of the discussion in this topic, consider Schrodinger's equation as given by

$$\hat{H}|\psi\rangle = i\hbar\frac{\partial}{\partial t}|\psi\rangle$$

A formal integration leads to

$$|\psi(t)\rangle = \sum_{\{n_i\}} e^{\hat{H}t/i\hbar}|\{n_i\}\rangle \qquad (4.2.6)$$

where

$$\hat{H} = \sum_k \hbar\omega_k\left(\hat{N}_k + \frac{1}{2}\right)$$

and where $\{n_i\}$ symbolizes the photons distributed across all the modes "i". Applying $\hat{H}$ to the Fock state in Equation 4.2.6 provides

$$|\psi(t)\rangle = \sum_{\{n_i\}} \exp\left[-\sum_k it\omega_k\left(n_k + \frac{1}{2}\right)\right]|n_1, n_2, \ldots\rangle$$

Now to discuss the solutions in more detail. As just mentioned, the Hamiltonian for a system of free-space photons is given by

$$\hat{H} = \sum_{\bar{k}s} \hbar\omega_k\left(\hat{b}^+_{\bar{k}s}\hat{b}_{\bar{k}s} + \frac{1}{2}\right) = \sum_{\bar{k}s} \hbar\omega_k\left(\hat{N}_{\bar{k}s} + \frac{1}{2}\right)$$

where the creation and annihilation operators depend on time according to

$$\hat{b}_{ks} = \hat{b}_{ks}(t) = \hat{b}_{ks}(0)e^{-i\omega_k t} \quad \text{and} \quad \hat{b}^+_{ks} = \hat{b}^+_{ks}(t) = \hat{b}^+_{ks}(0)e^{+i\omega_k t}$$

and satisfy the commutation relations

$$\left[\hat{b}_{\bar{k}s}, \hat{b}^+_{\bar{K}S}\right] = \delta_{\bar{k}\bar{K}}\delta_{sS} \quad \text{and} \quad \left[\hat{b}_{\bar{k}s}, \hat{b}_{\bar{K}S}\right] = 0 = \left[\hat{b}^+_{\bar{k}s}, \hat{b}^+_{\bar{K}S}\right]$$

In the following, the polarization index is suppressed. The number operator

$$\hat{N}_{\vec{k}} = \hat{b}^{+}_{\vec{k}}\hat{b}_{\vec{k}}$$

gives the number of photons with a particular wave vector $\vec{k}$ and polarization $\tilde{e}_{\vec{k}}$. Fock states are eigenvectors of the number operator according to

$$\begin{aligned}\hat{N}_{\vec{k}}\left|n_1,\ldots,n_{\vec{k}},\ldots\right\rangle &= \hat{b}^{+}_{\vec{k}}\hat{b}_{\vec{k}}\left|n_1,\ldots,n_{\vec{k}},\ldots\right\rangle \\ &= \hat{b}^{+}_{\vec{k}}\sqrt{n_{\vec{k}}}\left|n_1,\ldots,n_{\vec{k}}-1,\ldots\right\rangle \\ &= n_{\vec{k}}\left|n_1,\ldots,n_{\vec{k}},\ldots\right\rangle\end{aligned}$$

Therefore, the Fock states are seen to be eigenfunctions of the electromagnetic quantized Hamiltonian

$$\hat{H}\left|n_1,\ldots,n_{\vec{k}},\ldots\right\rangle = \sum_{\vec{K}}\hbar\omega_K\left(\hat{N}_{\vec{K}}+\frac{1}{2}\right)\left|n_1,\ldots,n_{\vec{k}},\ldots\right\rangle = \sum_{\vec{K}}\hbar\omega_K\left(n_{\vec{K}}+\frac{1}{2}\right)\left|n_1,\ldots,n_{\vec{k}},\ldots\right\rangle$$

For each basis state $\left|n_1,\ldots,n_{\vec{k}},\ldots\right\rangle$, the energy eigenvalue is

$$\sum_{\vec{k}}\hbar\omega_k\left(n_{\vec{k}}+\frac{1}{2}\right)$$

There is a different eigenvalue for each set of occupation numbers $n_1,n_2\ldots$.

Example: What is the energy eigenvalue corresponding to a single photon in the first mode of a Fabry Perot cavity? Assume the distance between the mirrors is L.

Solution: The Fock states is $\left|1,0,0\ldots\right\rangle$ and so we find

$$\hat{H}\left|1,0,0\ldots\right\rangle = \sum_{\vec{K}}\hbar\omega_K\left(\hat{N}_{\vec{K}}+\frac{1}{2}\right)\left|1,0,0\ldots\right\rangle = \hbar\omega_1\left(n_1+\frac{1}{2}\right)\left|1,0,0\ldots\right\rangle = \frac{3}{2}\hbar\omega_1\left|1,0,0\ldots\right\rangle$$

We can substitute for the angular frequency by writing $\omega = ck$ where

$$k = \frac{2\pi}{\lambda} = \frac{2\pi}{2L} = \frac{\pi}{L} \qquad \text{for the first } \textit{mode}.$$

So the total energy in the first mode is $E = \frac{3}{2}\hbar\omega_1 = \frac{3}{2}\hbar c\frac{\pi}{L}$

The energy stored in the Fock state $\left|n_1,\ldots,n_{\vec{k}},\ldots\right\rangle$ is given by

$$E = \sum_{\vec{k}}\hbar\omega_k\left(n_{\vec{k}}+\frac{1}{2}\right)$$

For the vacuum state $\left|0\right\rangle = \left|0,0,0,\ldots\right\rangle$, the stored energy (per volume) is

$$E = \sum_{\vec{k}}\frac{1}{2}\hbar\omega_k$$

The energy stored in the vacuum is infinite but there are no available quanta of energy. This energy corresponds to randomly oscillating electromagnetic fields that permeate all space (the vacuum fields). These fields are responsible for initiating spontaneous emission from an ensemble of excited atoms. There are theories that attribute vacuum effects to virtual photons created in the fields of charges and atoms so that the infinite vacuum energy can be eliminated.

Topic 4.2.3: Fluctuations of the EM Fields in Fock States

Returning to Equation 3.5.5 for the quantized electric field (repeated in Equation 4.2.7), the expectation value and standard deviation of the electric field for a Fock state is calculated. For simplicity, consider a single mode "k" traveling along the z-direction with frequency ω. Assume the available mode is the first one. The Fock state is then

$$|n\rangle \equiv |n,0,0\ldots\rangle$$

which describes the specific EM state. The electric field operator is given by

$$\hat{E} = +i\sum_{\bar{k}}\sqrt{\frac{\hbar\omega_k}{2\varepsilon_o V}}\left[\hat{b}_k\, e^{ikz-i\omega_k t} - \hat{b}_k^+\, e^{-ikz+i\omega_k t}\right]\tilde{e}_k \qquad (4.2.7)$$

which contains a sum over all possible modes.

First calculating the expectation value of the electric field in the Fock state with "n" photons gives

$$\langle n|\hat{E}|n\rangle = +i\sum_{\bar{k}}\sqrt{\frac{\hbar\omega_k}{2\varepsilon_o V}}\langle n,0,0\ldots|\left[\hat{b}_k\, e^{ikz-i\omega_k t} - \hat{b}_k^+\, e^{-ikz+i\omega_k t}\right]|n,0,0\ldots\rangle\tilde{e}_k$$

The summation over the modes can be eliminated since only one of them has any photons. Therefore

$$\langle n|\hat{E}|n\rangle = +i\sqrt{\frac{\hbar\omega}{2\varepsilon_o V}}\langle n|\left[\hat{b}\;\; e^{ikz-i\omega t} - \hat{b}^+\, e^{-ikz+i\omega t}\right]|n\rangle\tilde{e}$$

where k and ω are the wave vector and angular frequency of the mode.

$$\langle n|\hat{E}|n\rangle = +i\sqrt{\frac{\hbar\omega}{2\varepsilon_o V}}\langle n|\hat{b}|n\rangle e^{i\bar{k}\cdot\bar{r}-i\omega_k t}\,\tilde{e} - i\sqrt{\frac{\hbar\omega}{2\varepsilon_o V}}\langle n|\hat{b}^+|n\rangle e^{-i\bar{k}\cdot\bar{r}+i\omega_k t}\,\tilde{e}$$

$$= i\sqrt{\frac{\hbar\omega}{2\varepsilon_o V}}\sqrt{n}\langle n|n-1\rangle e^{i\bar{k}\cdot\bar{r}-i\omega_k t}\,\tilde{e} - i\sqrt{\frac{\hbar\omega}{2\varepsilon_o V}}\sqrt{n+1}\langle n|n+1\rangle e^{-i\bar{k}\cdot\bar{r}+i\omega_k t}\,\tilde{e}$$

Using the orthonormality of the Fock states

$$\langle n|m\rangle = \delta_{nm}$$

the *average* electric field is seen to be

$$\langle n|\hat{E}|n\rangle = 0$$

The average electric field in the Fock state is 0 because even though the photon number is exact for Fock state, the phase is completely unspecified. The idea is somewhat equivalent to integrating over the entire cycle of the sine wave. The fact that the photon number is exactly specified by the Fock state means that the phase (the conjugate variable) is completely unspecified in order to satisfy the number-phase Heisenberg uncertainty relation.

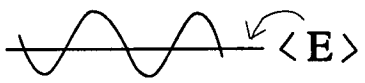


Figure 4.2.5: Representation of the average field.

Next, calculate the standard deviation of the electric field in the Fock state. Using the result for the average electric field, the variance is given by

$$\sigma_E^2 = \langle n|\left(\hat{E}^2 - \langle\hat{E}\rangle^2\right)|n\rangle = \langle n|\hat{E}^2|n\rangle$$

Using 4.2.7 and leaving off the summation, the square of the electric field operator is

$$\hat{E}^2 = -\frac{\hbar\omega}{2\varepsilon_o V}\left[\hat{b}\ e^{ikz-i\omega t} - \hat{b}^+ e^{-ikz+i\omega t}\right]^2$$

$$= -\frac{\hbar\omega}{2\varepsilon_o V}\left[\hat{b}^2 e^{2ikz-2i\omega t} + \left(\hat{b}^+\right)^2 e^{-2ikz+2i\omega t} - \hat{b}\hat{b}^+ - \hat{b}^+\hat{b}\right]$$

Now using the fact that $\langle n|\hat{b}^2|n\rangle = 0$ and $\langle n|\left(\hat{b}^+\right)^2|n\rangle = 0$,

$$\sigma_E^2 = \langle n|\hat{E}^2|n\rangle = -\frac{\hbar\omega}{2\varepsilon_o V}\langle n|\left(-\hat{b}\hat{b}^+ - \hat{b}^+\hat{b}\right)|n\rangle$$

Now use the commutation relation $\left[\hat{b},\hat{b}^+\right]=1$ so that

$$\hat{b}\hat{b}^+ = 1 + \hat{b}^+\hat{b}$$

can be substituted into the variance to get

$$\sigma_E^2 = \langle n|\hat{E}^2|n\rangle = -\frac{\hbar\omega}{2\varepsilon_o V}\langle n|\left(-1 - 2\hat{b}^+\hat{b}\right)|n\rangle = \frac{\hbar\omega}{\varepsilon_o V}\left(n + \frac{1}{2}\right) \qquad (4.2.8)$$

The last equation shows that even though the *average* electric field is zero for Fock states, the variance is never zero. Especially note that for the vacuum state with n=0, the standard deviation is

$$\sigma_E^2 = \langle n = 0|\hat{E}^2|n = 0\rangle = \frac{\hbar\omega}{2\varepsilon_o V} \qquad (4.2.9)$$

This is somewhat equivalent to integrating the square of electric field over a cycle. Equations 4.2.9 shows that there are electric field *fluctuations* in the vacuum and therefore the electric field at any given point in time is nonzero even though the average is zero. These vacuum field fluctuations initiate spontaneous emission from an ensemble of excited atoms. These fluctuations are equivalent to the zero point motion of the harmonic oscillator.

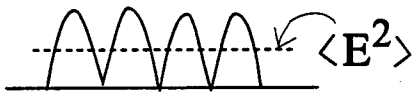


Figure 4.2.6: Representation of the variance of the electric field.

Section 4.3: Introduction to EM Coherent States

This section discusses coherent states of the electromagnetic field and contrasts them with Fock states. Although this section discusses the application to light, the reader should realize that it is applicable to RF electromagnetic energy, phonons and any other system that can be represented by a sum of "harmonic oscillators." Glauber and Yuen are the main contributors to the field of optical coherent states although the study extends back to the time of Schrodinger. The primary difference between classical light (large photon numbers) and the quantum mechanical counterpart is that, for each mode in the quantum mechanical description, the electric and magnetic field operators are the sum of two non-commuting terms. Consequently, the two terms cannot be simultaneously determined with infinite precision which leads to a nonzero variance for the electric field. The coherent state is discussed in general terms in this introductory section. The coherent state is related to the vacuum state by a translation of the vacuum to a state of non-zero average amplitude. A subsequent section discusses the mathematical foundation of coherent states and the stochastic models.

Topic 4.3.1: The Electric Field in the Coherent State

The light existing in the coherent state is the closest quantum mechanical analog to the classical picture of light as a sinusoidal wave. For the coherent state, the *average* amplitude of the electric (or magnetic) field is well defined (contrary to the unintuitive zero-average found for Fock states). The number of photons in a beam is not fixed, but instead follows a Poisson distribution. The larger the number of photons, the more nearly is the coherent state similar to the classical state of light. One major distinction between the classical and coherent descriptions, is that the coherent state requires uncertainty in the amplitude and phase of wave (i.e., noise).

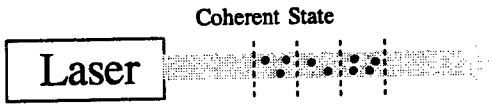


Figure 4.3.1: The number of photons in the coherent state follows a Poisson distribution.

The coherent state for a single optical mode is denoted by

$$|\alpha\rangle$$

where α is a complex number

$$\alpha = |\alpha| e^{i\phi} \qquad (4.3.1)$$

that is related to the *average* amplitude of the electric field and, most importantly, is an eigenvector of the annihilation operator

$$\hat{b}|\alpha\rangle = \alpha|\alpha\rangle$$

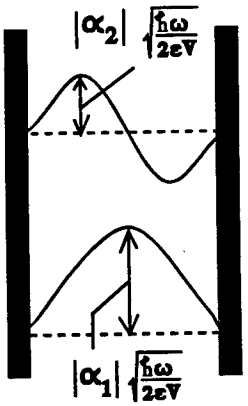


Figure 4.3.2: Modulus of the coherent state parameter α describes the amplitude of the field.

Until discussing the reason for calling α the *average* amplitude, it can be pictured as *the* amplitude (to within a multiplicative constant) as shown in Figure 4.3.2. For a system such as a Fabry-Perot cavity or a travelling wave with multiple modes, the coherent state is written as

$$|\alpha_1, \alpha_2, \ldots\rangle = |\alpha_1\rangle|\alpha_2\rangle\cdots$$

The multimode coherent state is actually a vector in a direct product space; each individual mode evolves independently of the other unless there is an explicit interaction between them (mitigated by an interaction Hamiltonian). The multimode coherent state is an eigenstate of the annihilation operator according to

$$\hat{b}_k|\alpha_1,\alpha_2,\ldots,\alpha_k,\ldots\rangle = \alpha_k|\alpha_1,\alpha_2,\ldots,\alpha_k,\ldots\rangle$$

The complex α_k is related to the average wave amplitude of mode #k. The reader will understand the definition of the coherent-state ket more fully as we continue. One important point is that two coherent states are *not* orthogonal even though we can normalize them to 1.

$$\langle\alpha|\beta\rangle \neq 0 \qquad\qquad \langle\alpha|\alpha\rangle = 1$$

The states are approximately orthogonal so long as α and β are sufficiently different (as discussed later). The vacuum state is the same for both Fock space and the space of coherent states.

Topic 4.3.2: Average Electric Field in the Coherent State

To understand the relation between the electromagnetic field amplitude and the coherent state, consider the expression for the quantized electric field found in Section 3.5 for the travelling waves

$$\hat{E} = +i\sum_{\vec{k}}\sqrt{\frac{\hbar\omega_k}{2\varepsilon_o V}}\left[\hat{b}_k\, e^{i\vec{k}\cdot\vec{r}-i\omega_k t} - \hat{b}_k^+\, e^{-i\vec{k}\cdot\vec{r}+i\omega_k t}\right]\tilde{e}_k \tag{4.3.2}$$

For simplicity, consider a single travelling mode "k" so that

$$\hat{E}_k = +i\sqrt{\frac{\hbar\omega_k}{2\varepsilon_o V}}\left[\hat{b}_k\, e^{i\vec{k}\cdot\vec{r}-i\omega_k t} - \hat{b}_k^+\, e^{-i\vec{k}\cdot\vec{r}+i\omega_k t}\right]\tilde{e}_k \tag{4.3.3}$$

Now suppose we calculate the average electric field in the state $|\alpha_k\rangle$. Noting that

$$\hat{b}_k|\alpha_k\rangle = \alpha_k|\alpha_k\rangle \quad\leftrightarrow\quad \langle\alpha_k|\hat{b}_k^+ = \langle\alpha_k|\,\alpha_k^* \tag{4.3.4}$$

by definition of the adjoint operator,

$$\begin{aligned}\langle\hat{E}_k\rangle = \langle\alpha_k|\hat{E}_k|\alpha_k\rangle &= +i\tilde{e}_k\sqrt{\frac{\hbar\omega_k}{2\varepsilon_o V}}\,\langle\alpha_k|\left\{\hat{b}_k\, e^{i\vec{k}\cdot\vec{r}-i\omega_k t} - \hat{b}_k^+\, e^{-i\vec{k}\cdot\vec{r}+i\omega_k t}\right\}|\alpha_k\rangle \\ &= +i\tilde{e}_k\sqrt{\frac{\hbar\omega_k}{2\varepsilon_o V}}\left\{\langle\alpha_k|\hat{b}_k|\alpha_k\rangle e^{i\vec{k}\cdot\vec{r}-i\omega_k t} - \langle\alpha_k|\hat{b}_k^+|\alpha_k\rangle e^{-i\vec{k}\cdot\vec{r}+i\omega_k t}\right\} \\ &= +i\tilde{e}_k\sqrt{\frac{\hbar\omega_k}{2\varepsilon_o V}}\left\{\alpha_k\, e^{i\vec{k}\cdot\vec{r}-i\omega_k t} - \alpha_k^*\, e^{-i\vec{k}\cdot\vec{r}+i\omega_k t}\right\}\end{aligned}$$

Using the expression found in Equation 4.3.1, $\alpha_k = |\alpha_k|\, e^{i\phi_k}$, the average field can be rewritten as

$$\langle\hat{E}_k\rangle = \langle\alpha_k|\hat{E}_k|\alpha_k\rangle = +i\tilde{e}_k\sqrt{\frac{\hbar\omega_k}{2\varepsilon_o V}}\left\{|\alpha_k|e^{i\vec{k}\cdot\vec{r}-i\omega_k t+i\phi_k} - |\alpha_k|e^{-i\vec{k}\cdot\vec{r}+i\omega_k t-i\phi_k}\right\}$$

Factor out the modulus and include the imaginary "i" with the exponentials to get

$$\left\langle \hat{E}_k \right\rangle = \langle \alpha_k | \hat{E}_k | \alpha_k \rangle = -\tilde{e}_k \sqrt{\frac{\hbar\omega_k}{2\varepsilon_o V}} \; |\alpha_k| \, \mathrm{Sin}\left(\vec{k}\cdot\vec{r} - \omega_k t + \phi_k\right)$$
$$= \tilde{e}_k \sqrt{\frac{\hbar\omega_k}{2\varepsilon_o V}} \; |\alpha_k| \, \mathrm{Sin}\left(\vec{k}\cdot\vec{r} - \omega_k t + \phi_k - \pi\right) \qquad (4.3.5)$$

Obviously, the α that appears in the ket for the coherent state $|\alpha\rangle$ is the phasor amplitude of the electric field (to within a normalization constant). It is possible to choose any desired phase for the average field just by adjusting ϕ_k. A later section will discuss the unitary operator that can be used to directly rotate the electric field operator rather than rotating the coherent state.

Topic 4.3.3: Coherent State as a Displaced Vacuum

The electric field operator $\hat{E}_k$ is a function of creation $\hat{b}_k^+$ and annihilation $\hat{b}_k$ operators

$$\hat{E}_k = +i \sqrt{\frac{\hbar\omega_k}{2\varepsilon_o V}} \left[\hat{b}_k \, e^{i\vec{k}\cdot\vec{r} - i\omega_k t} - \hat{b}_k^+ \, e^{-i\vec{k}\cdot\vec{r} + i\omega_k t}\right] \tilde{e}_k \qquad (4.3.6)$$

The two components of $\hat{E}_k$ cannot be simultaneous measured with infinite repeatability since the two amplitude operators do not commute as discussed in Chapter 3. Typically, the electric field is expressed in terms of quadratures using a new set of position $\hat{Q}_k$ and momentum $\hat{P}_k$ operators (for simplicity). These new electric field operator

$$\vec{E}_k(\vec{r},t) = -\tilde{e}_k \sqrt{\frac{\hbar\omega_k}{\varepsilon_o V}} \left[\hat{Q}_k \sin\left(\vec{k}\cdot\vec{r} - \omega_k t\right) + \hat{P}_k \cos\left(\vec{k}\cdot\vec{r} - \omega_k t\right)\right] \qquad (4.3.7)$$

obtains form the old one in Equation 4.3.7 by substituting the new operators

$$\hat{Q}_k = \frac{\omega_k \hat{q}_k}{\sqrt{\hbar\omega_k}} = \frac{1}{\sqrt{2}}\left[\hat{b}_k + \hat{b}_k^+\right] \qquad \hat{P}_k = \frac{\hat{p}_k}{\sqrt{\hbar\omega_k}} = -i\frac{1}{\sqrt{2}}\left[\hat{b}_k - \hat{b}_k^+\right]$$

The new "position and momentum" operators satisfy the new commutation relations

$$\left[\hat{Q}_k, \hat{P}_K\right] = \frac{\omega_k}{\hbar\omega_k}\left[\hat{q}_k, \hat{p}_K\right] = i\delta_{kK} \qquad \left[\hat{Q}_k, \hat{Q}_K\right] = 0 = \left[\hat{P}_k, \hat{P}_K\right]$$

Clearly, the two quadrature components for the same mode do not commute $\left[\hat{Q}_k, \hat{P}_k\right] = i$ which yields an uncertainty relation as discussed in Chapter 2. *Therefore, repeated measurements of the quadrature amplitdues* $\hat{Q}_k$ *and* $\hat{P}_k$ *yield a range of measured values* $\{Q_k\}$ *and* $\{P_k\}$, *respectively.* These ranges of values are discribed by a quasi-classical probability density (refer to the Wigner function).

The vacuum state (i.e., zero photon state) corresponds to a coherent state with zero average amplitude for the electric (and magnetic) field $\left\langle \hat{E}_k \right\rangle = \langle 0_k | \hat{E}_k | 0_k \rangle = 0$ as is easy to verify by setting α=0 in Equation 4.3.5. Section 7.5 shows that the electric field has a variance of

$$\sigma_E^2 = \langle 0|\hat{E}_k|0\rangle = \frac{\hbar\omega_k}{2\varepsilon_o V}$$

in the vacuum state. The variance (or equivalently, the standard deviation $\sigma_E = \Delta E_k$) is nonzero because the "position" $\hat{Q}_k$ and "momentum" $\hat{P}_k$ operators (or, equivalently, the creation $\hat{b}_k^\dagger$ and annihilation $\hat{b}_k$ operators) in the field operator $\hat{E}_k$ do not commute. However, we have also seen from Equation 4.2.4 that the wave function $u_o(Q_k)$ and $u_o(P_k)$ are Gaussian distributions although the wavefunctions were originally written in terms of $\hat{q}$ and $\hat{p}$. The vacuum wavefunctions are Gaussian in shape and provide the minimum uncertainty

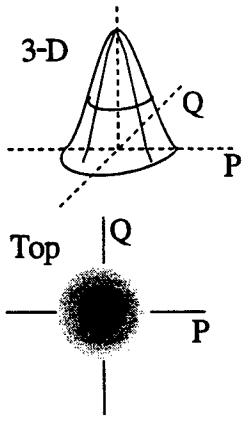


Figure 4.3.3: Quasi-classical probability distribution for the electric field in the vacuum state.

$$\Delta Q_k \, \Delta P_k = \left(\frac{\omega_k}{\sqrt{\hbar\omega_k}}\Delta q_k\right)\left(\frac{\Delta p_k}{\sqrt{\hbar\omega_k}}\right) = \frac{\omega_k}{\hbar\omega_k}\Delta q_k \Delta p_k = \frac{1}{2}$$

where, for the vacuum state, $\Delta q_k \Delta p_k = \hbar/2$. *Repeated* measurements of the quadratures produce a range of measured values Q_k and P_k (not the absence of the caret above the symbol). For the vacuum state, these values have an average of zero. The figure shows an example Wigner distribution for the possible values Q_k and P_k. The vacuum fluctuations (i.e., the standard deviation of P and Q) are represented by the radius of the distribution. A measurement of the electric field will produce a value somewhere within the circle.

Now, here's the main point. A coherent state is a displaced vacuum state (as we will demonstrate later). The parameter α in the coherent state $|\alpha\rangle$ is a complex number that gives the center of the distribution according to

$$\alpha_k = |\alpha_k| e^{i\phi} = \mathrm{Re}(\alpha_k) + i\,\mathrm{Im}(\alpha_k) = \frac{1}{\sqrt{2}}(Q_o + iP_o)$$

which is easy to verify by calculating

$$\langle \alpha_k|\hat{E}_k|\alpha_k\rangle$$

using Equation 4.3.7 with the definitions

$$\langle \alpha|\hat{Q}|\alpha\rangle = Q_o \qquad \langle \alpha|\hat{P}|\alpha\rangle = P_o$$

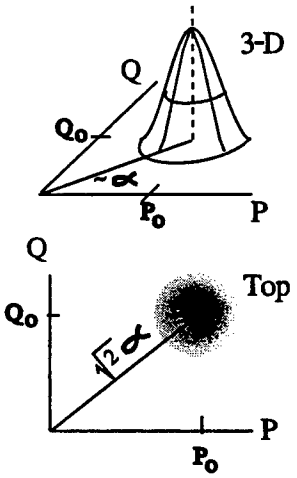


Figure 4.3.4: The coherent state is a displace vacuum.

Now it is clear why α represents the average amplitude. Notice that because the distribution has been translated without a change of shape, the amount of noise in the coherent state is exactly the same as the amount of noise in the vacuum.

Topic 4.3.4: The Nature of Quantum Noise

The term "noise" refers to the dispersion (i.e., standard deviation) in the electric field and quadrature terms. Each time a measurement is made of the electric field in the coherent state $|\alpha\rangle$, we expect to find a different value of amplitude and phase, denoted by the phasor α'. The amplitude and phase of each measured α' are both uncertain and are defined through the measured values Q and P, which define the domain of the Wigner distribution.

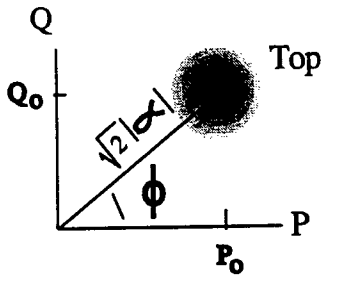


Figure 4.3.5: The amplitude and phase of the Wigner distribution.

$$\alpha'_k = |\alpha'_k|\, e^{i\phi'} = \mathrm{Re}(\alpha'_k) + i\,\mathrm{Im}(\alpha'_k) = \frac{1}{\sqrt{2}}(Q + iP)$$

An alternate way to express this is to say that the distance $\sqrt{2}\,|\alpha'_k|$ (which defines the amplitude) must be positioned within the circle representing the possible range of values. Likewise the phase ϕ' must position the radial vector within the boundaries of the circle. Therefore, each point α' or P,Q gives a different amplitude and phase of a sine wave. The measured waves can all differ by a phase and amplitude as shown in the figure. As the last piece of the "picture puzzle", again consider the top view of the distribution as shown in Figure 4.3.6.. The electric field operator $\hat{E}_k$ depends on time. In fact, the distribution of Q,P traces out a circle in the P-Q plot. The position of the a point in the circle corresponds to the particular amplitude and phase of the sinusoidal wave.

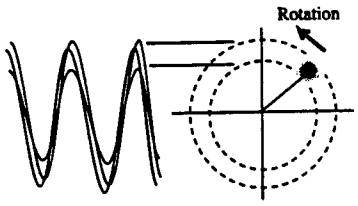


Figure 4.3.6: How the motion of the Wigner distribution provides a range of sine waves.

The "phase-space" plots (i.e., P-Q plots) show why two coherent states are only approximately orthogonal as indicated in Figure 4.3.7. The argument of the coherent-state ket (for example, α in $|\alpha\rangle$) represents the average amplitude. There can be significant overlap of two distributions so that an integral over phase space Q-P for the innerproduct $\langle\alpha|\beta\rangle$ (for example) is non-zero. However, two state widely separated in phase space have essentially zero overlap and the innerproduct $\langle\gamma|\delta\rangle \cong 0$. Apparently, as long as the two circles do not touch, the two corresponding coherent states will be approximately orthogonal. This is easy to see since the distribution for γ is zero where the distribution for δ is nonzero and vice versa so that the integral (for the inner product) is always zero.

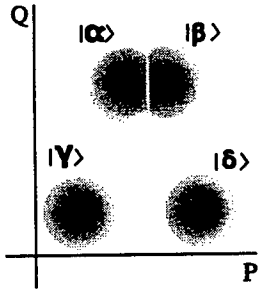


Figure 4.3.7: The overlap of coherent states control the inner product.

As a note, the structure of quantum theory regarding the relation between the operator and the state should be more evident now. The operators such as the Hamiltonian for the free field

$$\hat{H} = \sum_k \hbar\omega_k \left(\hat{b}_k^{+}\hat{b}_k + \frac{1}{2}\right)$$

always are written in the same way (no real need to find a new formula) and contain all the possible modes in the sum. The states describe the specifics of the system. The operator such as $\hat{H}$ can be used with either Fock states or coherent states. The expectation value of $\hat{H}$ in the Fock state $|n_1,0,\ldots\rangle$, for example, is

$$\langle n_1,0,\ldots|\hat{H}|n_1,0,\ldots\rangle = \hbar\omega_1\left(n_1+\frac{1}{2}\right)$$

whereas for the coherent state, using the same Hamiltonian, the expectation value is

$$\langle \alpha_1,0,\ldots|\hat{H}|\alpha_1,0,\ldots\rangle = \hbar\omega_1\left(|\alpha|^2+\frac{1}{2}\right)$$

Section 4.4: Definition and Statistics of Coherent States

The coherent state is the eigenvector of the annihilation operator. It is defined in as a sum over all Fock states. Unlike the Fock state, the number of photons in a coherent state follows a Poisson probability distribution and therefore does not have a well defined number of photons. The average and standard deviation (and higher moments) characterize the probability distribution for the photon number. The results are discussed in terms of the Signal-to-Noise ratio for a communications system.

Topic 4.4.1: The Coherent State from Fock States

The coherent state is defined to be the eigenstate of the annihilation operator according to

$$\hat{b}_k|\alpha_1,\ldots,\alpha_k,\ldots\rangle = \alpha_k\,|\alpha_1,\ldots,\alpha_k,\ldots\rangle$$

The complex parameter $\alpha_k = |\alpha_k|\,e^{i\phi_k}$ is the phasor representation of the electric field amplitude (to within a normalization constant) for mode "k". Obviously, the coherent state $|\alpha_1,\ldots,\alpha_k,\ldots\rangle = |\alpha_1\rangle|\alpha_2\rangle\ldots$ is a vector in a direct product space. For simplicity, focus on a single mode $|\alpha_k\rangle$ and therefore drop the subscript "k". The basic definition of the coherent state is then

$$\hat{b}|\alpha\rangle = \alpha\,|\alpha\rangle \tag{4.4.1}$$

By applying the adjoint operator to both sides of Equation 4.4.1, the basic definition can be equivalently stated as

$$\langle\alpha|b^+ = \langle\alpha|\alpha^*$$

The following discussion shows that the coherent state is given by

$$|\alpha\rangle = e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle \tag{4.4.2}$$

where the Fock state $|n\rangle = |n_k\rangle$ keeps track of the number of photons in mode "k". Recall that $\{|n_k\rangle\}$ spans a single-mode space that might actually be part of a multimode space (i.e., direct product space) of the form $|n_1,\ldots,n_k,\ldots\rangle = |n_1\rangle|n_2\rangle\ldots|n_k\rangle\ldots$. Equation 4.4.2 is quite easy to work with.

Now for the reasoning behind Equation 4.4.2. We are looking for eigenstates of the annihilation operator. Try a linear combination of Fock states.

$$|\alpha\rangle = \sum_{n=0}^{\infty}C_n|n\rangle \tag{4.4.3}$$

Apply the annihilation operator to Equation 4.4.3 and require Equation 4.4.1 to hold

$$\alpha|\alpha\rangle = \hat{b}|\alpha\rangle = \sum_{n=0}^{\infty}C_n\hat{b}|n\rangle = \sum_{n=0}^{\infty}C_n\sqrt{n}\,|n-1\rangle$$

Substitute Equation 4.4.3 for the first term in the previous expression to obtain

$$\sum_{n=0}^{\infty}\alpha C_n|n\rangle = \sum_{n=1}^{\infty}C_n\sqrt{n}\,|n-1\rangle$$

where the second sum starts at n=1 since $\sqrt{0}=0$. A recusion relation can be found for the expansion coefficients C_n. In the second summation, let $n-1 \rightarrow n$ to get

$$\sum_{n=0}^{\infty} \alpha C_n |n\rangle = \sum_{n=0}^{\infty} C_{n+1} \sqrt{n+1}\, |n\rangle$$

Comparing sides (or equivalently, operating with $\langle m|$ on both sides) provides

$$C_{n+1} = C_n \frac{\alpha}{\sqrt{n+1}}$$

Assume C_0 is known. The following sequence results

$$C_0 \qquad C_1 = C_0 \frac{\alpha}{\sqrt{1}} \qquad C_2 = C_1 \frac{\alpha}{\sqrt{2}} = C_0 \frac{\alpha^2}{\sqrt{1 \cdot 2}} \qquad \ldots \qquad C_n = C_0 \frac{\alpha^n}{\sqrt{n!}}$$

Now Equation 4.4.3 can be rewritten as

$$|\alpha\rangle = \sum_{n=0}^{\infty} C_0 \frac{\alpha^n}{\sqrt{n!}} |n\rangle \tag{4.4.4}$$

Normalizing the coherent state vector to 1 yields the constant C_0

$$1 = \langle \alpha | \alpha \rangle = \left(\sum_{m=0}^{\infty} C_0 \frac{\alpha^m}{\sqrt{m!}} |m\rangle \right)^{+} \sum_{n=0}^{\infty} C_0 \frac{\alpha^n}{\sqrt{n!}} |n\rangle = \sum_{mn} C_0^* C_0 \frac{(\alpha^m)^*}{\sqrt{m!}} \frac{\alpha^n}{\sqrt{n!}} \langle m|n\rangle$$

Using the orthonormality relation for Fock states

$$\langle m|n\rangle = \delta_{mn}$$

provides

$$1 = \langle \alpha | \alpha \rangle = |C_0|^2 \sum_n \frac{|\alpha|^{2n}}{n!}$$

Comparing this last expression with the Taylor series expansion of

$$e^x = 1 + \frac{x}{1} + \frac{x^2}{2!} + \ldots = \sum_n \frac{x^n}{n!}$$

gives

$$\sum_n \frac{|\alpha|^{2n}}{n!} = \exp |\alpha|^2$$

and the constant C_0 is

$$1 = |C_0|^2 \exp |\alpha|^2 \quad \rightarrow \quad C_0 = e^{-|\alpha|^2/2}$$

where the phase is set equal to unity. Finally, Equations 4.4.3 and 4.4.4 can be written as

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \tag{4.4.5}$$

Topic 4.4.2: The Poisson Distribution

This topic derives the Poisson probability distribution that characterizes the photon number in a coherent state. The coherent state is therefore shown to exhibit shot noise.

What is the probability that there are "m" photons in the electromagnetic coherent state $|\alpha\rangle$? The question is answered by using the fact that the coherent state resides in a vector space spanned by the single-mode Fock states $\{|n\rangle\}$ by virtue of Equation 4.4.5

$$|\alpha\rangle = e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle \qquad (4.4.6)$$

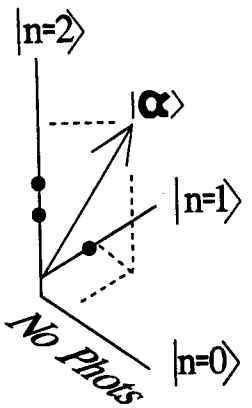


Figure 4.4.1: The coherent state as an element of Fock space by virtue of Eq. 4.4.6. The solid circles indicate the number of photons residing in the corresponding Fock state.

The probability *amplitude* for the coherent state having "m" photons is found by projecting the coherent state $|\alpha\rangle$ onto the basis state $|m\rangle$ (which is the Fock state with the number of photons "m"). Therefore, the *probability* of m-photons in coherent state $|\alpha\rangle$ is

$$P_\alpha(m) = |\langle m|\alpha\rangle|^2 \qquad (4.4.7)$$

Recall that the quantity $\langle m|\alpha\rangle$ is an expansion coefficient similar to those discussed in Chapter 2.

The probability $P_\alpha(m)$ can be found as follows. First operate on Equation 4.4.6 with $\langle m|$ to get

$$\langle m|\alpha\rangle = e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}\langle m|n\rangle = e^{-|\alpha|^2/2}\frac{\alpha^m}{\sqrt{m!}}$$

since $\langle m|n\rangle = \delta_{mn}$. Substituting into Equation 4.4.7 provides

$$P_\alpha(m) = |\langle m|\alpha\rangle|^2 = e^{-|\alpha|^2}\frac{|\alpha|^{2m}}{m!} \qquad (4.4.8)$$

This expression can be compared with the Poisson probability distribution usually written as

$$P(m) = \frac{\xi^m e^{-\xi}}{m!} \qquad (4.4.9)$$

where $\xi = |\alpha|^2$ is the *average number of photons* in the coherent state $|\alpha\rangle$. There is a possibility of having extremely large numbers of photons in the beam even for small field amplitudes α. As a note, Equations 4.4.8 and 4.4.9 are probabilities and must sum to unity according to

$$\sum_m P_\alpha(m) = \sum_m |\langle m|\alpha\rangle|^2 = \sum_m e^{-|\alpha|^2}\frac{|\alpha|^{2m}}{m!} = 1 \qquad (4.4.10)$$

Topic 4.4.3: Average Photon Number

What is the average number of photons $\langle m \rangle$ in a field characterized by the coherent state $|\alpha\rangle$? The following discussion shows that the average number is

$$\bar{n} = \xi = |\alpha|^2$$

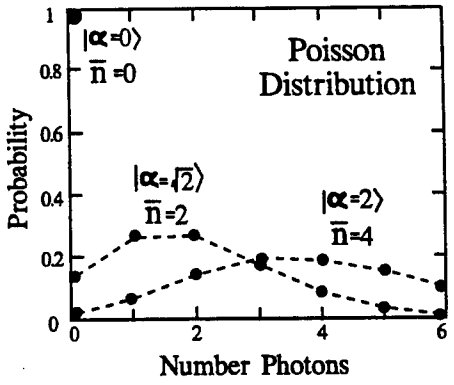


Figure 4.4.2: The Poisson distribution for averages of 0, 2, 4 photons in a mode.

Figure 4.4.2 shows the discrete Poisson distribution for three coherent states. There is only one point for the coherent state $|\alpha = 0\rangle$ since it is also the Fock vacuum state $|0\rangle$ which definitely has no photons (i.e., the standard deviation is zero). The average occupation number is shown by two methods. The first method uses classical probability theory and the Poisson distribution while the second method uses operators.

Method 1: Classical Probability Theory

The average of "m" is calculated according to

$$\langle m \rangle = \sum_{m=0}^{\infty} m P_\alpha(m)$$

Substituting Equation 4.4.9 provides

$$\langle m \rangle = \sum_{m=0}^{\infty} m \frac{\xi^m e^{-\xi}}{m!} = \sum_{m=1}^{\infty} m \frac{\xi^m e^{-\xi}}{m!} = \sum_{m=1}^{\infty} \frac{\xi^m e^{-\xi}}{(m-1)!}$$

Let $m-1 \rightarrow m$ to get

$$\langle m \rangle = \sum_{m=0}^{\infty} \frac{\xi^{m+1} e^{-\xi}}{m!} = \xi \sum_{m=0}^{\infty} \frac{\xi^m e^{-\xi}}{m!} = \xi$$

where the final step obtains from the fact that the probabilities add to 1 as in Equation 4.4.10. So the average number of photons in the coherent state $|\alpha\rangle$ is

$$\langle m \rangle = \xi = |\alpha|^2$$

Method 2: Operators

Let $\hat{N} = \hat{b}^+ \hat{b}$ be the number operator. The expectation number of photons is then

$$\bar{n} \equiv \langle \hat{N} \rangle = \langle \alpha | \hat{b}^+ \hat{b} | \alpha \rangle = \left[\hat{b} | \alpha \rangle \right]^+ \hat{b} | \alpha \rangle = \left[\alpha | \alpha \rangle \right]^+ \alpha | \alpha \rangle = |\alpha|^2 \langle \alpha | \alpha \rangle = |\alpha|^2$$

since $|\alpha\rangle$ is an eigenstate of the annihilation operator.

Topic 4.4.4: Standard Deviation

What is the standard deviation for the number of photons in an electromagnetic mode characterized by the coherent state $|\alpha\rangle$? Recall that the standard deviation σ_N is found from the variance according to

$$\sigma_N^2 = \langle \hat{N}^2 \rangle - \langle \hat{N} \rangle^2 = \langle \alpha | \hat{N}^2 | \alpha \rangle - \langle \alpha | \hat{N} | \alpha \rangle^2 \qquad (4.4.11)$$

The variance could be calculated similar to the classical Method 1 used to calculate the average. However, the operator approach is simpler. The last term in Equation 4.4.11 is already known from the previous topic. Now calculate the first term.

$$\langle \alpha | \hat{N}^2 | \alpha \rangle = \langle \alpha | \hat{b}^+ \hat{b}\, \hat{b}^+ \hat{b} | \alpha \rangle$$

The middle two operators need to be commuted using the commutation relations

$$\left[\hat{b}, \hat{b}^+\right] = 1 \rightarrow \hat{b}\hat{b}^+ = \hat{b}^+\hat{b} + 1$$

to get

$$\langle \alpha | \hat{N}^2 | \alpha \rangle = \langle \alpha | \hat{b}^+ \left(\hat{b}^+\hat{b} + 1\right)\hat{b} | \alpha \rangle = \langle \alpha | \hat{b}^+\hat{b}^+\hat{b}\hat{b} | \alpha \rangle + \langle \alpha | \hat{b}^+\hat{b} | \alpha \rangle$$

The rest is even easier. Use relations of the form

$$\hat{b}|\alpha\rangle = \alpha|\alpha\rangle \quad \rightarrow \quad \hat{b}\hat{b}|\alpha\rangle = \alpha^2|\alpha\rangle$$

(and so on) to find

$$\langle \alpha | \hat{N}^2 | \alpha \rangle = \langle \alpha | \hat{b}^+\hat{b}^+\hat{b}\hat{b} | \alpha \rangle + \langle \alpha | \hat{b}^+\hat{b} | \alpha \rangle = \left(\alpha^*\right)^2 \alpha^2 \langle \alpha | \alpha \rangle + \alpha^* \alpha \langle \alpha | \alpha \rangle$$

The coherent states are normalized to one, so that

$$\langle \alpha | \hat{N}^2 | \alpha \rangle = |\alpha|^4 + |\alpha|^2$$

Equation 4.4.11 provides the variance as

$$\sigma_N^2 = \langle \hat{N}^2 \rangle - \langle \hat{N} \rangle^2 = \langle \alpha | \hat{N}^2 | \alpha \rangle - \langle \alpha | \hat{N} | \alpha \rangle^2 = |\alpha|^4 + |\alpha|^2 - |\alpha|^4 = |\alpha|^2 = \bar{n}$$

so the standard deviation is

$$\sigma_N = \sqrt{\langle n \rangle}$$

Topic 4.4.5: Signal-to-Noise Ratio

The signal to noise ratio for a coherent beam of light generated by a laser (or other means) is important for communications. The signal strength varies as the average number of photons in the beam. The standard deviation is a measure of the noise. The Signal-to-Noise Ratio is

$$S/N = \frac{\langle n \rangle}{\sqrt{\langle n \rangle}} = \sqrt{\langle n \rangle}$$

For small numbers of photons, the inherent unavoidable noise of the light is a larger percentage of the signal. For example, an optical beam with 100 photons has a standard deviation of 10 and a signal-to-noise ratio of 10.

For systems composed of a small number of atoms, such as nanometer scale devices, low S/N can be a problem. For example, a device with dimensions smaller than

100x100x100 angstrom might consist of 30x30x30 atoms (using about 3 angstroms per atom). The total number of atoms must be less than 27,000. Now if the collection is electrically pumped so that 10% are emitting light at any particular time then there are at most 2700 photons. The standard deviation in this case is about 50. The expected total variation of the signal is roughly twice the standard deviation or about 100. Therefore, the signal can be expected to vary by at least 4% due to inherent quantum noise. The percentage can be higher for systems with fewer atoms. For many analogue applications, this is an unacceptably high noise level. Subsequent sections show that it might be possible to reduce detected noise by working with "squeezed states."

Topic 4.4.6: Uncertainty in Energy for Coherent States

The standard deviation of the energy for a mode can be calculated using

$$\hat{H}_k = \hbar\omega_k\left(N_k + \frac{1}{2}\right)$$

The standard deviation is

$$\sigma_H = \hbar\omega_k \; \sigma_N = \hbar\omega_k\sqrt{\bar{n}}$$

Section 4.5: Coherent States as Displaced Vacuum States

The introduction to coherent states (Section 4.3) discusses the coherent state as an eigenvector of the annihilation operator and its role as a displaced vacuum. The present section fills in the mathematical detail of the displacement operation and the phase space operators (P,Q). Any coherent state can be obtained by displacing the vacuum state in phase space (P,Q). For simplicity, restrict the discussion to a single optical mode. Previous sections show that the coherent state for a field with zero amplitude $|\alpha = 0\rangle$ is identical to the vacuum Fock state $|0\rangle$. The single mode vacuum state has a coordinate wave function with a Gaussian shape. Therefore, the coordinate wavefunction for the non-zero coherent state is expected to also be a Gaussian since the coherent state is a displaced vacuum. This section discusses the operator that performs the displacement and also derives the coordinate wavefunction.

Topic 4.5.1: The Displacement Operator

The definition of the coherent state as a sum over Fock states from Section 4.4 provides the displacement operator. The coherent state is

$$|\alpha\rangle = e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle \tag{4.5.1}$$

where $\{|n\rangle = |n_k\rangle\}$ are single mode Fock states and "n" is the photon occupation number. If the Fock state $|n\rangle$ in Equation 4.5.1 is related to the vacuum state $|0\rangle$, then the equation provides a transformation from the vacuum state to the coherent state $|\alpha\rangle$. The transformation between the two states is the displacement operator. To relate the Fock state to the vacuum, use the boson creation operators create the Fock states

$$|1\rangle = \frac{\hat{b}^+}{\sqrt{1}}|0\rangle \qquad |2\rangle = \frac{\hat{b}^+}{\sqrt{2}}|1\rangle = \frac{\left(\hat{b}^+\right)^2}{\sqrt{2\cdot 1}}|0\rangle \qquad \cdots \qquad |n\rangle = \frac{\left(\hat{b}^+\right)^n}{\sqrt{n!}}|0\rangle \tag{4.5.2}$$

Consequently, the coherent state in Equation 4.5.1 becomes

$$|\alpha\rangle = e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{\alpha^n\left(\hat{b}^+\right)^n}{n!}|0\rangle$$

However, the summation can be written as an exponential to give

$$|\alpha\rangle = e^{-|\alpha|^2/2}e^{\alpha\hat{b}^+}|0\rangle \tag{4.5.3}$$

Equation 4.5.3 shows explicitly that the coherent state $|\alpha\rangle$ is a displaced vacuum state. The "displacement operator" is

$$\hat{D}(\alpha) = e^{-|\alpha|^2/2}e^{\alpha\hat{b}^+}$$

which is a unitary operator (as shown later). It is customary to rewrite the displacement operator by using a Taylor expansion

$$e^{-\alpha^*\hat{b}}|0\rangle = \left[1 - \alpha^*\hat{b} + \frac{\left(\alpha^*\hat{b}\right)^2}{2!} + \cdots\right]|0\rangle = |0\rangle$$

(where $\hat{b}^n|0\rangle = 0$) which provides the typical form of the displacement operator

$$|\alpha\rangle = e^{-|\alpha|^2/2} e^{\alpha\hat{b}^+} e^{-\alpha^*\hat{b}}|0\rangle \tag{4.5.4}$$

In some cases it is convenient to combine the three exponentials into a single one. Chapter 3 discusses the commutation relations that the arguments of the exponentials must satisfy in order to combine the exponentials. In particular, the Campbell-Baker-Hausdorff equation states

$$\exp\left(\hat{A}+\hat{B}\right) = \exp\hat{A}\exp\hat{B}\exp\left[-\frac{[\hat{A},\hat{B}]}{2}\right]$$

so long as $\left[\hat{A},\left[\hat{A},\hat{B}\right]\right] = 0 = \left[\hat{B},\left[\hat{A},\hat{B}\right]\right]$. Setting $\hat{A} = \alpha\hat{b}^+$ and $\hat{B} = -\alpha^*\hat{b}$, provides

$$D(\alpha) = e^{\alpha\hat{b}^+ - \alpha^*\hat{b}} \tag{4.5.5}$$

and

$$|\alpha\rangle = e^{\alpha\hat{b}^+ - \alpha^*\hat{b}}|0\rangle \tag{4.5.6}$$

Topic 4.5.2: Properties of the Displacement Operator

(1) The displacement operator is unitary with $\hat{D}^+(\alpha) = \hat{D}^{-1}(\alpha) = \hat{D}(-\alpha)$

The displacement operator is unitary. As discussed in Chapter 2, an operator $u = e^{i\hat{O}}$ is unitary so long as $\hat{O}$ is Hermitian since

$$\hat{u}\hat{u}^+ = e^{i\hat{O}}e^{-i\hat{O}} = e^{i\hat{O}}e^{-i\hat{O}} = 1$$

The argument in the displacement operator $D(\alpha) = e^{\alpha\hat{b}^+ - \alpha^*\hat{b}}$ can be written as $i\hat{O} = \alpha\hat{b}^+ - \alpha^*\hat{b}$ where $\hat{O} = -i(\alpha\hat{b}^+ - \alpha^*\hat{b})$ is easily seen to be Hermitian.

The inverse of D is easily seen to be

$$D(-\alpha) = e^{-(\alpha\hat{b}^+ - \alpha^*\hat{b})}$$

since then

$$D(\alpha)\,D(-\alpha) = e^{\alpha\hat{b}^+ - \alpha^*\hat{b}}e^{-(\alpha\hat{b}^+ - \alpha^*\hat{b})} = e^{\alpha\hat{b}^+ - \alpha^*\hat{b} - (\alpha\hat{b}^+ - \alpha^*\hat{b})} = 1$$

where the exponentials were combined because the arguments $i\hat{O} = \alpha\hat{b}^+ - \alpha^*\hat{b}$ and $-i\hat{O}$ commute $\left[i\hat{O}, -i\hat{O}\right] = 0$.

(2) The displaced creation and annihilation operators are found by a similarity transformation

$$D^+(\alpha)\hat{b}^+ D(\alpha) = \hat{b}^+ + \alpha \qquad D^+(\alpha)\hat{b}D(\alpha) = \hat{b} + \alpha$$

For example, consider the second relation. The operator expansion theorem from Chapter 2, which is

$$e^{-\hat{A}}\hat{B}e^{\hat{A}} = \hat{B} - [\hat{A},\hat{B}] + \frac{1}{2!}\left[\hat{A},[\hat{A},\hat{B}]\right] + \ldots$$

with $\hat{A} = \alpha\hat{b}^+ - \alpha^*\hat{b}$ and $\hat{B} = \hat{b}$ provides

$$e^{-(\alpha\hat{b}^+-\alpha^*\hat{b})}\hat{b}\,e^{\alpha\hat{b}^+-\alpha^*\hat{b}} = \hat{b}-\left[\alpha\hat{b}^+-\alpha^*\hat{b},\hat{b}\right]+0+\ldots=\hat{b}+\alpha$$

(3) The displacement operator acting on a nonzero coherent state produces another coherent state with the sum of two amplitudes and a complex phase factor.

(4) The "phase space" representation of the displacement operator.

The "phase space" operators $\hat{Q},\hat{P}$ are defined through

$$\hat{b}=\frac{\omega\hat{q}}{\sqrt{2\hbar\omega}}+\frac{i\hat{p}}{\sqrt{2\hbar\omega}}=\frac{\hat{Q}}{\sqrt{2}}+\frac{i\hat{P}}{\sqrt{2}}$$

and

$$\hat{b}^+=\frac{\omega\hat{q}}{\sqrt{2\hbar\omega}}-\frac{i\hat{p}}{\sqrt{2\hbar\omega}}=\frac{\hat{Q}}{\sqrt{2}}-\frac{i\hat{P}}{\sqrt{2}}$$

where

$$\left[\hat{Q},\hat{P}\right]=\left[\frac{\omega\hat{q}}{\sqrt{\hbar\omega}},\frac{\hat{p}}{\sqrt{\hbar\omega}}\right]=\frac{\omega}{\hbar\omega}[\hat{q},\hat{p}]=i$$

The center of the Wigner distribution is defined to be at Q_o,P_o with the phasor α in the coherent state $|\alpha\rangle$ given by

$$\alpha=\frac{1}{\sqrt{2}}\left[P_o+iQ_o\right]$$

Therefore, the displacement operator can be written

$$D(\alpha)=\exp\left(\alpha\hat{b}^+-\alpha^*\hat{b}\right)=\exp\left[\frac{\alpha}{\sqrt{2}}\left(\hat{Q}-i\hat{P}\right)-\frac{\alpha^*}{\sqrt{2}}\left(\hat{Q}+i\hat{P}\right)\right]=\exp\left[iP_o\hat{Q}-iQ_o\hat{P}\right. \qquad (4.5.7)$$

The derivation of the Wigner function makes use of Equation 4.5.7. An alternate form of this equation is useful for finding the coordinate representation of the coherent state (similar to $\psi(q)$ for Fock space). The Campbell-Baker-Hausdorff equation from Chapter 2

$$\exp\left(\hat{A}+\hat{B}\right)=\exp\hat{A}\exp\hat{B}\exp\left[-\frac{[\hat{A},\hat{B}]}{2}\right] \qquad \left[\hat{A},\left[\hat{A},\hat{B}\right]\right]=0=\left[\hat{B},\left[\hat{A},\hat{B}\right]\right]$$

with $\hat{A}=iP_o\hat{Q}$ and $\hat{B}=-iQ_o\hat{P}$ yields $\left[\hat{A},\hat{B}\right]=\left[iP_o\hat{Q},-iQ_o\hat{P}\right]=(i)(-i)P_oQ_o\left[\hat{Q},\hat{P}\right]=iP_oQ_o$ and

$$D(\alpha)=\exp\left[iP_o\hat{Q}-iQ_o\hat{P}\right]=\exp\left(iP_o\hat{Q}\right)\exp\left(-iQ_o\hat{P}\right)\exp\left(-\frac{i}{2}P_oQ_o\right) \qquad (4.5.8)$$

Topic 4.5.3: The Coordinate Representation of a Coherent State

Let $|\alpha\rangle$ be a single-mode coherent state, which is a vector in an abstract Hilbert space. The objective of this section is to represent the coherent state as a function of the "position" coordinate, denoted by $U_\alpha(Q)=\langle Q|\alpha\rangle$, rather than as a vector. This is similar to the coordinate wave functions found for the Fock states. Although the vector is nice to show the vacuum displacement, it does not explicitly show the range of electric field amplitudes Q to be expected. $U_\alpha(Q)$ is a useful representation of the coherent state because it shows that the "electric field amplitudes" Q are normally distributed. In addition, the functions $U_\alpha(Q)$ pave the conceptual path for the Wigner distribution (refer to Figure 4.5.1 below). This section presents two methods of proof that the fields are normally distributed. The first method makes use of the basic definition of the coherent state as an eigenstate of the annihilation operator. The second method demonstrates the use of the displacement operator.

Method 1: Coordinate Representation of the Coherent State using the Annihilation Operator

Method 1 to obtain the coordinate representation of the coherent state $|\alpha\rangle$ treats the state as an eigenvector of the annihilation operator and then solves a first order differential equation. The method emphasizes the simplest method without resorting to the use of the displacement operator.

By definition, the arbitrary coherent state $|\alpha\rangle$ with the complex amplitude $\alpha=|\alpha|\,e^{i\phi}$ is an eigenstate of the annihilation operator

$$\hat{b}|\alpha\rangle=\alpha|\alpha\rangle \tag{4.5.9}$$

The annihilation operator can be written in terms of the position and momentum operators

$$\hat{b}=\frac{\omega\hat{q}}{\sqrt{2\hbar\omega}}+\frac{\hat{p}}{\sqrt{2\hbar\omega}}=\frac{\hat{Q}}{\sqrt{2}}+\frac{i\hat{P}}{\sqrt{2}} \tag{4.5.10}$$

We can work with either the usual position $\hat{q}$ and momentum $\hat{p}$ operators or the new ones $\hat{Q},\hat{P}$ used for our "phase-space" representations. The new momentum operator has the "coordinate" representation

$$\hat{P}=\frac{\hat{p}}{\sqrt{\hbar\omega}}=\frac{1}{\sqrt{\hbar\omega}}\frac{\hbar}{i}\frac{\partial}{\partial q}=\frac{1}{\sqrt{\hbar\omega}}\frac{\hbar}{i}\frac{\partial Q}{\partial q}\frac{\partial}{\partial Q}=\frac{1}{i}\frac{\partial}{\partial Q} \tag{4.5.11}$$

where use was made of the relation from 4.5.10

$$Q=\frac{\omega q}{\sqrt{\hbar\omega}}$$

Alternatively, as a side note, it is possible to find the coordinate representation of the new momentum operator $\hat{P}$ by letting "c" be a "constant to be determined" in

$$\hat{P} = c\frac{\partial}{\partial Q}$$

and requiring the new momentum and position operators to satisfy their commutation relation

$$\left[\hat{Q},\hat{P}\right] = i \;\rightarrow\; \left[Q, c\frac{\partial}{\partial Q}\right] = i$$

Letting the second commutator operate on an arbitrary function f(Q) shows that the only choice for "c" is c=1/i which agrees with the results for 4.5.11 which is

$$\hat{P} = \frac{1}{i}\frac{\partial}{\partial Q}$$

Working in the Q coordinate representation, project both sides of Equation 4.5.9 onto the coordinate Q to obtain

$$\langle Q|\, \hat{b}\left(\hat{Q},\hat{P}\right)|\alpha\rangle = \alpha\langle Q|\alpha\rangle \qquad (4.5.12a)$$

where the notation $\hat{b}\left(\hat{Q},\hat{P}\right)$ serves as a reminder that the annihilation operator depends on the position and momentum operators before the coordinate projection. Define the coordinate representation of the coherent state $|\alpha\rangle$ to be $U_\alpha(Q) = \langle Q|\alpha\rangle$, which is a wavefunction. Equation 4.5.12 becomes

$$\hat{b}(Q)\, U_\alpha(Q) = \alpha\, U_\alpha(Q) \qquad (4.5.12b)$$

where now the annihilation operator depends on Q and the derivative with respect to Q

$$\hat{b}(\hat{Q},\hat{P}) = \frac{\hat{Q}}{\sqrt{2}} + \frac{i\hat{P}}{\sqrt{2}} \;\rightarrow\; \hat{b}(Q) = \frac{1}{\sqrt{2}}\left(Q + \frac{\partial}{\partial Q}\right)$$

Equation 4.5.12b can be written as

$$\frac{1}{\sqrt{2}}\left(Q + \frac{\partial}{\partial Q}\right)U_\alpha(Q) = \alpha U_\alpha(Q)$$

or

$$\left(Q - \alpha\sqrt{2} + \frac{\partial}{\partial Q}\right)U_\alpha(Q) = 0$$

This is a first order, ordinary differential equation with the solution

$$U_\alpha(Q) = C_\alpha \exp\left[-\frac{\left(Q-\alpha\sqrt{2}\right)^2}{2}\right] \qquad (4.5.13)$$

Normalizing the function U provides the constant C_α (by setting $\int_{-\infty}^{\infty} dQ\, U^* U = 1$).

$$C_\alpha = \frac{1}{\pi^{1/4}}\exp\left(\mathrm{Im}\,\alpha\right)^2 \qquad (4.5.14)$$

Equation 4.5.13 shows that the electric field amplitude (represented by Q) is normally distributed and centered at $\alpha\sqrt{2}$. Figure 4.5.1 shows that $U_\alpha(Q)$ is the projection of the Wigner distribution onto the plane containing the Q-

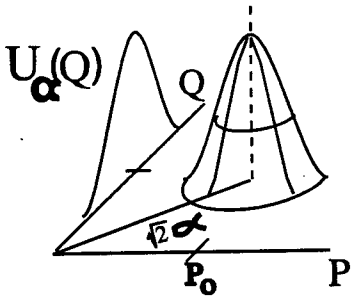


Figure 4.5.1: The Wigner distribution of the coherent state $|\alpha\rangle$ is shown as the 3-D relief plot. The coordinate representation $U_\alpha(Q)$ is the projection onto the plane containg the Q-axis.

axis. A momentum wave function $U_\alpha(P)$ would provide a similiar distribution for the plane containing the P-axis. Therefore $U_\alpha(Q)$ and $U_\alpha(P)$ are "shadows" from which to deduce the Wigner distribution.

Method 2: Coordinate Representation of the Coherent State using the Displacement Operator

Method 2 obtains the coordinate representation directly by displacing the vacuum state, which is known to have a Gaussian distribution. It emphasizes the notion of the displacement operator as a generator of translations in the abstract Hilbert space. The method does not require the solution of a differential equation but it is more involved than the first method.

The displacement operator translates the vacuum state $|0\rangle$ to the coherent state $|\alpha\rangle$ according to

$$|\alpha\rangle = D(\alpha,\hat{Q},\hat{P})|0\rangle \qquad (4.5.15)$$

where the operators appear explicitly in the argument of D. Equation 4.5.8 provides the appropriate form for the displacement operator

$$D(\alpha,\hat{Q},\hat{P}) = \exp\left[iP_o\hat{Q} - iQ_o\hat{P}\right] = \exp\left(iP_o\hat{Q}\right)\exp\left(-iQ_o\hat{P}\right)\exp\left(-\frac{i}{2}P_oQ_o\right) \qquad (4.5.16)$$

where

$$\alpha = \frac{1}{\sqrt{2}}(Q_o + iP_o)$$

Projecting the Equations 4.5.15 and 4.5.16 onto the Q coordinates provides

$$U_\alpha(Q) = D(\alpha,Q)U_0(Q) \qquad (4.5.17)$$

In Equation 4.5.17, $U_\alpha(Q)$ is the Q-coordinate representation of the coherent state α, that is $\langle Q|\alpha\rangle = U_\alpha(Q)$. The operator $D(\alpha,Q)$ obtains from Equation 4.5.15 using

$$\hat{Q} \to Q \qquad \text{and} \qquad \hat{P} \to \frac{1}{i}\frac{\partial}{\partial Q}$$

since the procedure is to be carried out in the Q-coordinate representation. The last term in Equation 4.5.16 is a complex constant. The middle term requires some discussion.

When Equation 4.5.16 is combine with Equation 4.5.17, one of the factors has the form

$$\text{factor} = \exp\left(-iQ_o\hat{P}\right)U_0(Q) \qquad (4.5.17)$$

To realize that the exponential is a translation operator in Q-space consider a Taylor series expansion of the function $U_0(Q+\xi_i)$ about the point Q (where ξ_k is a small addition to Q). The expansion provides

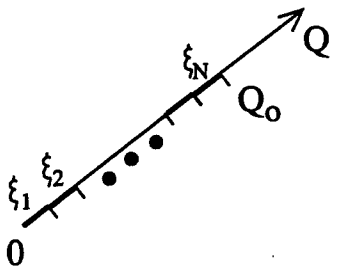


Figure 4.5.2: The displacement Q_o is made of smaller displacements ξ_i.

$$U_o(Q+\xi_k) \cong U_o(Q) + \frac{\partial U_o(Q)}{\partial Q}\xi_k + \ldots = \left(1 + \xi_k \frac{\partial}{\partial Q} + \ldots\right) U_o(Q)$$

Replacing the derivative with

$$\hat{P} \leftrightarrow \frac{1}{i}\frac{\partial}{\partial Q}$$

gives

$$U_o(Q+\xi_k) = \left(1 + \xi_k \frac{\partial}{\partial Q} + \ldots\right) U_o(Q) = \left(1 + i\xi_k \hat{P} + \ldots\right) U_o(Q) = \exp\left(+i\xi_k \hat{P}\right) U_o(Q)$$

Now, by repeatedly applying the infinitesimal translation operator, we can build up the entire length Q_o

$$U_o(Q+Q_o) = \prod_k \exp\left(i\xi_k \hat{P}\right) U_o(Q) = \exp\left(\sum_k i\xi_k \hat{P}\right) U_o(Q) = \exp\left(iQ_o \hat{P}\right) U_o(Q)$$

So the exponential with the momentum operator is a translation. Replacing Q_o with $-Q_o$ gives the effect of the exponential in the "factor" in Equation 4.5.17.

$$\text{factor} = \exp\left(-iQ_o \hat{P}\right) U_0(Q) = U_0(Q - Q_o)$$

Continuing to work with the combinitation of Equations 4.5.16 and 4.5.17

$$U_\alpha(Q) = D(\alpha, Q) U_0(Q) = \exp(iP_o Q) \exp\left(-iQ_o \hat{P}\right) \exp\left(-\frac{i}{2} P_o Q_o\right) U_0(Q)$$

$$= \exp(iP_o Q) \left\{\exp\left(-iQ_o \hat{P}\right) U_0(Q)\right\} \exp\left(-\frac{i}{2} P_o Q_o\right)$$

$$= \exp(iP_o Q)\, U_0(Q - Q_o) \exp\left(-\frac{i}{2} P_o Q_o\right)$$

Substituting the expression for U_o

$$U_0(Q) = \frac{1}{\pi^{1/4}} \exp\left[-\frac{Q^2}{2}\right]$$

into the last equation provides

$$U_\alpha(Q) = \frac{1}{\pi^{1/4}} \exp\left[-\frac{(Q-Q_o)^2}{2} + iP_o Q - \frac{iP_o Q_o}{2}\right]$$

which agrees with the *combination* of Equations 4.5.13 and 4.5.14.

As an important note, the coordinate space wavefunctions are never specified as U(P,Q) since the phase space operators P,Q do not commute and cannot be consistently specified together. However, the Wigner function provides a semiclassical probability distribution for which it is possible to speak of P,Q together.

Section 4.6: Quasi-Orthonormality, Closure and Trace for Coherent States

Previous sections for the coherent state shows how it is produced from the vacuum state under the action of the displacement operator. They also show the coordinate representation of the coherent state in anticipation of the Wigner distribution. The present section verifies more of the claims made in the introduction regarding properties such as quasi-orthonormality, completeness, closure and the trace operation.

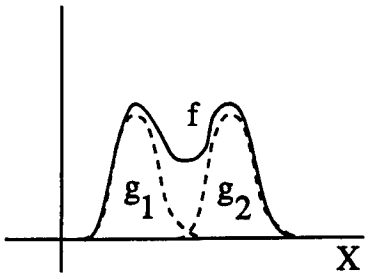


Figure 4.6.1: The function "f" is the sum of two different Gaussians g_1 and g_2.

Each coherent state resides in an abstract Hilbert space. However the collection of coherent states is not a vector space! The summation over Fock states used to define the coherent state, allows only a subset of the vectors in Fock space. In fact, the summation induces the property that all coherent states are Gaussians. A function such as "f" in Figure 4.6.1, which is the sum of two arbitrary Gaussians g_1 and g_2, is not necessary itself a Gaussian. Therefore, the property of closure for a vector space is violated. If the collection of coherent states are considered to be basis vectors for another space, then the vector space properties hold; however, the set of "basis" vectors (i.e., the coherent states) is over-complete.

Topic 4.6.1: Normalization

The single mode coherent states are defined by

$$|\alpha\rangle = e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle \qquad \text{or equivalently} \qquad \langle\alpha| = e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{(\alpha^*)^n}{\sqrt{n!}}\langle n|$$

They are normalized as can be seen

$$\langle\alpha|\alpha\rangle = e^{-|\alpha|^2}\sum_{mn=0}^{\infty}\frac{(\alpha^*)^m(\alpha)^n}{\sqrt{m!}\sqrt{n!}}\langle m|n\rangle = e^{-|\alpha|^2}\sum_{mn=0}^{\infty}\frac{(\alpha^*)^m(\alpha)^n}{\sqrt{m!}\sqrt{n!}}\delta_{mn} = e^{-|\alpha|^2}\sum_{n=0}^{\infty}\frac{|\alpha|^{2n}}{n!} = 1$$

since

$$\sum_{n=0}^{\infty}\frac{|\alpha|^{2n}}{n!} = e^{|\alpha|^2}$$

Topic 4.6.2: Quasi-Orthogonality

Let $|\alpha\rangle$ and $|\beta\rangle$ be two coherent states.

$$\langle\alpha|\beta\rangle = \exp\left[-\frac{|\alpha|^2+|\beta|^2}{2}\right]\sum_{n,m=0}^{\infty}\frac{(\alpha^*)^m}{\sqrt{m!}}\frac{\beta^n}{\sqrt{n!}}\langle m|n\rangle$$

$$= \exp\left[-\frac{|\alpha|^2+|\beta|^2}{2}\right]\sum_{n,m=0}^{\infty}\frac{(\alpha^*\beta)^n}{n!}$$

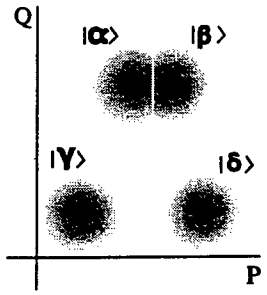


Figure 4.6.2: The overlap of coherent states control the inner product.

where the orthonormality of the single-mode Fock states $\langle m|n\rangle = \delta_{mn}$ was used. The summation gives an exponential.

$$\langle\alpha|\beta\rangle = \exp\left[-\frac{|\alpha|^2+|\beta|^2}{2}\right]\exp\left(\alpha^*\beta\right) = \exp\left[-|\alpha-\beta|^2\right] \tag{4.6.1}$$

The overlap between the two states decreases exponentially as the coherent states are separated. This behavior is consistent with the fact that the Wigner probability distributions are Gaussian. Equation 4.6.1 shows that when $\alpha=\beta$, the inner product is one.

Example: What is the inner product between the coherent state $|\alpha\rangle$ with an average of $\bar{n}_\alpha = 25$ photons and $|\beta\rangle$ having $\bar{n}_\beta = 16$. Assume that α,β are real.

Solution: Section 4.4 shows that $|\alpha|^2 = \bar{n}$. Ignoring the phase, the amplitudes are $\alpha = \sqrt{\bar{n}_\alpha} = 5$ and $\beta = 4$. Therefore $\langle\alpha|\beta\rangle = \exp\left[-(5-4)^2\right] = e^{-1} = 0.37$. If one of the states has nonzero phase, then the states are further separated and the overlap is negligible.

Topic 4.6.3: Closure

The set of single-mode coherent states $\{|\alpha\rangle\}$ satisfy a closure relation of the form

$$\frac{1}{\pi}\int_{\substack{\alpha-\\ \text{plane}}} d^2\alpha\ |\alpha\rangle\langle\alpha| = 1 \tag{4.6.2}$$

where α is a complex number $\alpha = re^{i\phi} = \alpha_x + i\alpha_y$ and the integral is over the entire α-plane with $d^2\alpha = d\alpha_x\, d\alpha_y$. The set of coherent states form a basis but they are over-complete (i.e., not all of them are needed to span the space).

Substituting the Fock expansion for the coherent state begins the proof of Equation 4.6.2

$$|\alpha\rangle = e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle = e^{-r^2/2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle$$

to get

$$\frac{1}{\pi}\int_{\substack{\alpha-\\ \text{plane}}} d^2\alpha\ |\alpha\rangle\langle\alpha| = \frac{1}{\pi}\sum_{n,m}\frac{|n\rangle\langle m|}{\sqrt{n!m!}}\int d^2\alpha\, e^{-r^2/2}\left(\alpha^*\right)^n e^{-r^2/2}\alpha^m = \frac{1}{\pi}\sum_{n,m}\frac{|n\rangle\langle m|}{\sqrt{n!m!}}\int d^2\alpha\, e^{-r^2}\alpha^{*n}\alpha^m$$

Substituting the polar-coordinate element of area $d^2\alpha$ and writing α in polar form provides

$$\frac{1}{\pi}\int_{\substack{\alpha-\\ \text{plane}}} d^2\alpha\ |\alpha\rangle\langle\alpha| = \frac{1}{\pi}\sum_{n,m}\frac{|n\rangle\langle m|}{\sqrt{n!m!}}\int r\,dr\,d\phi\ e^{-r^2} r^{m+n} e^{i(m-n)\phi} = \frac{1}{\pi}\sum_{n,m}\frac{|n\rangle\langle m|}{\sqrt{n!m!}}\int dr\,d\phi\ e^{-r^2} r^{m+n+1} e^{i(m-n)\phi}$$

The integral over the angle provides

$$\int_0^{2\pi} d\phi\, e^{i(m-n)\phi} = 2\pi\delta_{mn}$$

since for $m \neq n$ the integral is over a multiple number of complete cycles. The closure integral becomes

$$\frac{1}{\pi}\int_{\substack{\alpha-\\ \text{plane}}} d^2\alpha\ |\alpha\rangle\langle\alpha| = 2\sum_{n,m}\frac{|n\rangle\langle m|}{\sqrt{n!m!}}\delta_{mn}\int dr\ e^{-r^2} r^{m+n+1} = 2\sum_{n}\frac{|n\rangle\langle n|}{n!}\int_0^{\infty} dr\ e^{-r^2} r^{2n+1}$$

Integral tables provide the last integral

$$\int_0^{\infty} dr\ e^{-r^2} r^{2n+1} = \frac{n!}{2}$$

Therefore the closure integral becomes

$$\frac{1}{\pi}\int_{\substack{\alpha-\\ \text{plane}}} d^2\alpha\ |\alpha\rangle\langle\alpha| = 2\sum_{n}\frac{|n\rangle\langle n|}{n!}\frac{n!}{2} = \sum_{n}|n\rangle\langle n| = 1$$

as required.

Topic 4.6.4: Coherent-State Expansion of a Fock State

The objective is to express the single mode Fock state $|n\rangle$ as an expansion of coherent states. This is accomplished by using the coherent-state closure relation

$$|n\rangle = 1|n\rangle = \left(\frac{1}{\pi}\int d^2\alpha\ |\alpha\rangle\langle\alpha|\right)|n\rangle = \frac{1}{\pi}\int d^2\alpha\ |\alpha\rangle\langle\alpha|n\rangle$$

$\langle\alpha|n\rangle$ is a probability amplitude which is related to the Poisson probability distribution. Evaluate the inner product using the definition of coherent state gives

$$\langle\alpha|n\rangle = e^{-|\alpha|^2/2}\sum_{m=0}^{\infty}\frac{(\alpha^*)^m}{\sqrt{m!}}\langle m|n\rangle = e^{-|\alpha|^2/2}\sum_{m=0}^{\infty}\frac{(\alpha^*)^m}{\sqrt{m!}}\delta_{mn} = e^{-|\alpha|^2/2}\frac{(\alpha^*)^n}{\sqrt{n!}}$$

so the Fock state $|n\rangle$ is

$$|n\rangle = 1|n\rangle = \frac{1}{\pi}\int d^2\alpha\ |\alpha\rangle\langle\alpha|n\rangle = \int d^2\alpha\ |\alpha\rangle\ e^{-|\alpha|^2/2}\frac{(\alpha^*)^n}{\pi\sqrt{n!}}$$

Notice that each Fock basis vector is a linear combination of the coherent states; consequently, every vector in Fock space is a linear combination of the coherent states. The coherent states span Fock space.

Topic 4.6.5: Over-Completeness of Coherent States

The set of vectors $\{|\alpha\rangle\}$ is over-complete in the sense that each one can be expressed as a sum over the others. The situation is similar to having three vectors span a 2-D vector space; obviously, one of them is not required.

The fact that one coherent-state vector can be expressed as a sum of the others can be seen as follows

$$|\alpha\rangle = 1|\alpha\rangle = \frac{1}{\pi}\int d^2\beta\ |\beta\rangle\langle\beta|\alpha\rangle$$

Using the inner product between two coherent states given in Equation 4.6.1

$$\langle\beta|\alpha\rangle = \exp\left[-|\beta-\alpha|^2\right]$$

provides

$$|\alpha\rangle = \frac{1}{\pi}\int d^2\beta\ |\beta\rangle \exp\left[-|\beta-\alpha|^2\right]$$

The more separated are the parameters α and β the less the state $|\beta\rangle$ contributes to the state $|\alpha\rangle$. The integral is thought of as a summation.

Topic 4.6.6: Trace of an Operator Using Coherent States

The trace formula involves the factor of "$1/\pi$" similar to the closure relation. Starting with the definition of trace using single mode Fock states, then inserting the coherent-state closure relation, and then removing the Fock states using the Fock state closure relation can find the formula.

$$\mathrm{Tr}\hat{O} = \sum_n \langle n|\hat{O}|n\rangle = \sum_n \langle n|1\hat{O}|n\rangle = \sum_n \langle n|\left\{\frac{1}{\pi}\int d^2\alpha\ |\alpha\rangle\langle\alpha|\right\}\hat{O}|n\rangle = \frac{1}{\pi}\int d^2\alpha \sum_n \langle n|\alpha\rangle\langle\alpha|\hat{O}|n\rangle$$

Interchanging the order of the matrix elements in the last term gives

$$\mathrm{Tr}\hat{O} = \frac{1}{\pi}\int d^2\alpha \sum_n \langle\alpha|\hat{O}|n\rangle\langle n|\alpha\rangle = \frac{1}{\pi}\int d^2\alpha\ \langle\alpha|\hat{O}\left\{\sum_n |n\rangle\langle n|\right\}|\alpha\rangle = \frac{1}{\pi}\int d^2\alpha\ \langle\alpha|\hat{O}|\alpha\rangle$$

Therefore, the formula is similar to the trace formula using Fock states except that an integral appears along with the factor of $1/\pi$.

$$\mathrm{Tr}\hat{O} = \frac{1}{\pi}\int d^2\alpha\ \langle\alpha|\hat{O}|\alpha\rangle$$

Section 4.7: Field Fluctuations in the Coherent State

This section shows that electromagnetic fields exhibit minimum uncertainty for the coherent states. The uncertainty, in this case, is measured as the variance of the electromagnetic field (at any point in space-time). The electric field is shown to have smaller variance for coherent states than for Fock states. This is exactly opposite to the results for the Hamiltonian, which has minimum variance for the Fock states (the variance is zero).

For this section, recall that the (single mode) electric field can be written in either of the two equivalent forms as

$$\hat{E}_k = +i\sqrt{\frac{\hbar\omega}{2\varepsilon_o V}}\left[\hat{b}\, e^{i\vec{k}\cdot\vec{r}-i\omega t} - \hat{b}^+ e^{-i\vec{k}\cdot\vec{r}+i\omega t}\right] \tag{4.7.1a}$$

or

$$\vec{E}_k(\vec{r},t) = -\sqrt{\frac{\hbar\omega}{\varepsilon_o V}}\left[\hat{Q}\sin(\vec{k}\cdot\vec{r}-\omega t)+\hat{P}\cos(\vec{k}\cdot\vec{r}-\omega t)\right] \tag{4.7.1b}$$

where the mode subscripts and polarization vector are suppressed, and $\omega, \vec{k}$ are the angular frequency and wave-vector of the traveling waves. The creation $\hat{b}^+$ and annihilation $\hat{b}$ operatators are related to the "position" $\hat{Q}$ and "momentum" $\hat{P}$ operators according to

$$\hat{Q} = \frac{1}{\sqrt{2}}\left[\hat{b}+\hat{b}^+\right]$$

$$\hat{P} = \frac{-i}{\sqrt{2}}\left[\hat{b}-\hat{b}^+\right]$$

The operators satisfy the following commutation relations

$$\left[\hat{b},\hat{b}^+\right]=1 \qquad \left[\hat{b},\hat{b}\right]=0=\left[\hat{b},\hat{b}^+\right]$$

$$\left[\hat{Q},\hat{P}\right]=i \qquad \left[\hat{Q},\hat{Q}\right]=0=\left[\hat{P},\hat{P}\right]$$

As discussed in the Section 4.5, the coherent states have Gaussian probability distribution functions in the coordinate representation.

Gaussian wave functions produce the "minimum-area" uncertainty relation.

$$\Delta Q\,\Delta P = \frac{1}{2} \tag{4.7.2}$$

The term "area" is used because $\Delta Q\,\Delta P$ is an area in phase space. The term "minimum" indicates the use of the "=" sign; for any other wave function, the minimum area is larger than ½. Some developments of quantum statistical mechanics divide up phase space into small rectangles each with the "minimum area" (see, for example, Pathria). Similar ideas can be applied to quantum optics. The uncertainty relation indicates that the quadrature terms cannot simultaneously *and* precisely be measured owing to the non-commutivity of $\hat{Q}$ and $\hat{P}$. Recall from Section 2.11, that if Hermitian operators $\hat{A},\hat{B}$ satisfy $\left[\hat{A},\hat{B}\right]=i\hat{C}$ then $\sigma_a\sigma_b \geq \frac{1}{2}\left|\langle\hat{C}\rangle\right|$, which does not depend on the type of state vector (such as Fock,

coherent, squeezed). Therefore, regardless of the type of state vector, the Q-P uncertainty relation always has the form

$$\Delta Q\,\Delta P \geq \frac{1}{2}$$

In the ensuing topics, the uncertainty relation are determined for the quadrature components in the Fock and coherent states. As will be shown, coherent states produce minimum-area uncertainty relations and minimum variance electromagentic fields.

Topic 4.7.1: The Uncertainty Relation for Coherent States

The minimum-area uncertainty relation requires that the standard deviation for $\hat{Q}$ and $\hat{P}$ be evaluated using coherent states. The variance of $\hat{Q}$ is found from

$$\sigma_Q^2 = \langle\alpha|\hat{Q}^2|\alpha\rangle - \langle\alpha|\hat{Q}|\alpha\rangle^2$$

First calculate the average $\langle\alpha|\hat{Q}|\alpha\rangle$ by using the creation-annihilation operators for the "position" operator $\hat{Q}$ and also the definition of the coherent state

$$\hat{b}|\alpha\rangle = \alpha|\alpha\rangle \quad \leftrightarrow \quad \langle\alpha|\hat{b}^+ = \langle\alpha|\alpha^*$$

The expectation value is

$$\langle\alpha|\hat{Q}|\alpha\rangle^2 = \frac{1}{2}\langle\alpha|\left[\hat{b}+\hat{b}^+\right]|\alpha\rangle^2 = \frac{1}{2}\left(\alpha+\alpha^*\right)^2 = \frac{\alpha^2+\alpha^{*2}}{2}+|\alpha|^2$$

Next calculate $\langle\alpha|\hat{Q}^2|\alpha\rangle$ using the commutation relation $\left[\hat{b},\hat{b}^+\right]=1$

$$\langle\alpha|\hat{Q}^2|\alpha\rangle = \frac{1}{2}\langle\alpha|\left[\hat{b}+\hat{b}^+\right]^2|\alpha\rangle = \frac{1}{2}\langle\alpha|\left[\hat{b}^2+\left(\hat{b}^+\right)^2+\hat{b}\hat{b}^++\hat{b}^+\hat{b}\right]|\alpha\rangle$$

$$= \frac{1}{2}\left(\alpha^2+\alpha^{*2}\right)+\frac{1}{2}\langle\alpha|\left(\hat{b}\hat{b}^++\hat{b}^+\hat{b}\right)|\alpha\rangle = \frac{1}{2}\left(\alpha^2+\alpha^{*2}\right)+\frac{1}{2}\langle\alpha|\left(2\hat{b}^+\hat{b}+1\right)|\alpha\rangle$$

$$= \frac{\alpha^2+\alpha^{*2}}{2}+|\alpha|^2+\frac{1}{2}$$

Therefore, the variance of $\hat{Q}$ is

$$(\Delta Q)^2 = \sigma_Q^2 = \langle\alpha|\hat{Q}^2|\alpha\rangle - \langle\alpha|\hat{Q}|\alpha\rangle^2 = \frac{1}{2}$$

Similarly, we can show the variance of the "momentum" operator in the coherent state is

$$(\Delta P)^2 = \sigma_P^2 = \frac{1}{2}$$

Regardless of the value of α, the uncertainty relation for the coherent state is

$$\Delta Q\,\Delta P = \frac{1}{2} \tag{4.7.3}$$

The uncertainty relation for coherent states (4.7.3) could have alternatively been calculated using vacuum expectation values. The reason, as has been previously pointed out, is that the coherent state is a displaced vacuum; the noise is unaffected by the

displacement operation. This result is obvious from the derivation of Equation 4.7.3 by setting $\alpha=0$ and realizing that 4.7.3 is independent of α.

Topic 4.7.2: Field and Energy Variance for Coherent States

The variance for the electric field in the coherent state is

$$\sigma_E^2 = \langle\alpha|\hat{E}^2|\alpha\rangle - \langle\alpha|\hat{E}|\alpha\rangle^2 = -\frac{\hbar\omega}{2\varepsilon_o V}\left\{\langle\alpha|\left[\hat{b}\,e^{i\vec{k}\cdot\vec{r}-i\omega t} - \hat{b}^+ e^{-i\vec{k}\cdot\vec{r}+i\omega t}\right]^2|\alpha\rangle - \langle\alpha|\left[\hat{b}\,e^{i\vec{k}\cdot\vec{r}-i\omega t} - \hat{b}^+ e^{-i\vec{k}\cdot\vec{r}+i\omega t}\right]|\alpha\rangle^2\right\}$$

Being careful to use the commutation relations, the variance for the electric field evaluated in a coherent state is found to be

$$\sigma_E^2 = \frac{\hbar\omega}{2\varepsilon_o V}$$

Notice that the variance is independent of the value α. The coherent state represents the minimum achievable dispersion for the electric field. Repeated measurements of the electric field result in a range of numbers characterized by σ_E.

The variance of the energy in the coherent state can be calculated from the Hamiltonian

$$\hat{H} = \hbar\omega\left(\hat{b}^+\hat{b} + \frac{1}{2}\right)$$

The variance can now be calculated

$$\sigma_H^2 = \langle\alpha|\hat{H}^2|\alpha\rangle - \langle\alpha|\hat{H}|\alpha\rangle^2 = (\hbar\omega)^2\langle\alpha|\left(\hat{b}^+\hat{b} + \frac{1}{2}\right)^2|\alpha\rangle - (\hbar\omega)^2\langle\alpha|\left(\hat{b}^+\hat{b} + \frac{1}{2}\right)|\alpha\rangle^2$$

$$= (\hbar\omega)^2\langle\alpha|\left(\hat{b}^+\hat{b}\hat{b}^+\hat{b} + \hat{b}^+\hat{b} + \frac{1}{4}\right)|\alpha\rangle - (\hbar\omega)^2\langle\alpha|\left(\hat{b}^+\hat{b} + \frac{1}{2}\right)|\alpha\rangle^2$$

The first term on the right-hand side has $\hat{b}\hat{b}^+$; the order of these two operators must be commuted to evaluate the term.

$$\sigma_H^2 = (\hbar\omega)^2\langle\alpha|\left(\hat{b}^+\hat{b}^+\hat{b}\hat{b} + 2\hat{b}^+\hat{b} + \frac{1}{4}\right)|\alpha\rangle - (\hbar\omega)^2\langle\alpha|\left(\hat{b}^+\hat{b} + \frac{1}{2}\right)|\alpha\rangle^2$$

$$= (\hbar\omega)^2\left\{|\alpha|^4 + 2|\alpha|^2 + \frac{1}{4}\right\} - (\hbar\omega)^2\left(|\alpha|^4 + |\alpha|^2 + \frac{1}{4}\right)$$

$$= (\hbar\omega)^2|\alpha|^2$$

Recalling that the average number of photons in the coherent state $|\alpha\rangle$ is given by

$$\bar{n}_\alpha = |\alpha|^2$$

So, the energy in the coherent state is not definite. Repeated measurements of the energy produces a range of values characterized by the standard deviation

$$\Delta H = \sigma_H = \hbar\omega\sqrt{\bar{n}}$$

The uncertainty in the energy is due to the fluctuations in the electric field. The power in the electromagnetic beam is expected to fluctuate. In the vacuum state $\alpha=0$, there is not any uncertainty in the energy since the vacuum is also a Fock state.

Topic 4.7.3: Comparison of Variance for Coherent and Fock States

The Q-P uncertainty relation for single-mode Fock states can be calculated

$$\sigma_P^2 = \langle n|\hat{P}^2|n\rangle - \langle n|\hat{P}|n\rangle^2 = \langle n|\hat{P}^2|n\rangle = \langle n|\frac{-i}{\sqrt{2}}\left[\hat{b}-\hat{b}^+\right]^2|n\rangle = -\frac{1}{2}\langle n|\left(\hat{b}^2+\hat{b}^{+2}-\hat{b}^+\hat{b}-\hat{b}\hat{b}^+\right)|n\rangle$$

Noting $\langle n|\hat{b}^2|n\rangle \sim \langle n|n+2\rangle = 0$ with similar results for the creation operator and using the commutation relations provides

$$\sigma_P^2 = \langle n|\left(\hat{b}^+\hat{b}+\frac{1}{2}\right)|n\rangle = n+\frac{1}{2}$$

where use is made of the fact that $\hat{N} = \hat{b}^+\hat{b}$ is the number operator. A similar result holds for the "position" operator

$$\sigma_Q^2 = n+\frac{1}{2}$$

Therefore, the Q-P uncertainty relation for a Fock state $|n\rangle$

$$\Delta Q\,\Delta P = \sigma_Q\sigma_P = n+\frac{1}{2} \geq \frac{1}{2}$$

The "position-momentum" uncertainty is always larger that ½ which is the uncertainty in the coherent state.

$$(\Delta Q\,\Delta P)_{\text{Fock}} > (\Delta Q\,\Delta P)_{\text{coherent}} \qquad n>0 \quad (\text{non-vacuum})$$

$$(\Delta Q\,\Delta P)_{\text{Fock}} = (\Delta Q\,\Delta P)_{\text{coherent}} = \frac{1}{2} \qquad n=0 \quad (\text{vacuum})$$

The coherent states are the minimum uncertainty states. The noise for the electric field in a coherent state is called the "standard quantum limit" (SQL). Essentially it is the lowest possible noise level. Squeezing techniques can reduce the noise in one quadrature; however, the noise in the other increases (similarly for "number" and "phase"). This behavior occurs because the uncertainty relation

$$\Delta Q\,\Delta P \geq \frac{1}{2}$$

must still hold.

The uncertainty in the electric field evaluated in a Fock state is calculated in Topic 7.5.3

$$\sigma_E^2 = \frac{\hbar\omega}{\varepsilon_o V}\left(n+\frac{1}{2}\right) \qquad \text{Fock}$$

for the Fock state $|n\rangle$. Therefore, the uncertainty in the electric field is always larger for the Fock state than for the coherent state (except for the vacuum state).

Finally, given that Fock states are eigenstates of the Hamiltonian, the variance of the energy must be zero

$$\sigma_H = 0 \qquad \text{Fock}$$

The uncertainty in energy and power is always larger for the coherent states. However, Fock states are very difficult to achieve in practice. The number-squeezed state is the closest relative to the Fock state. All of the number-noise (i.e., amplitude noise) would need to be squeezed out of the coherent state to turn it into a Fock state. This shows the

reason for the *number-phase* uncertainty relation. Removing noise (i.e., reducing the standard deviation) from the amplitude necessarily requires the uncertainty in the phase to increase in order to maintain the Q-P uncertainty relations. Removing all of the number-noise causes the phase to be completely unspecified.

Section 4.8: Introduction to Squeezed States

The previous few sections discuss "classical" electromagnetic waves as the large amplitude limit of electromagnetic fields in coherent states. The coherent states produce the minimum uncertainty (i.e., minimum fluctuations) in the field. Even though Fock states produce sharp values for the energy and power, they cannot be realized in practice and therefore, practically speaking, the coherent states produce the minimum uncertainty in the electromagnetic energy and power. Electromagnetic waves are characterized by amplitude and phase, which cannot be simultaneously and precisely measured. Coherent states offer the best compromise. Electromagnetic waves are expressed in quantized form in terms of amplitudes or quadratures as

$$\hat{E} = +i\sqrt{\frac{\hbar\omega}{2\varepsilon_o V}}\left[\hat{b}\, e^{i\vec{k}\cdot\vec{r}-i\omega t} - \hat{b}^+ e^{-i\vec{k}\cdot\vec{r}+i\omega t}\right]$$

or

$$\vec{E}(\vec{r},t) = -\sqrt{\frac{\hbar\omega}{\varepsilon_o V}}\left[\hat{Q}\sin\left(\vec{k}\cdot\vec{r}-\omega t\right) + \hat{P}\cos\left(\vec{k}\cdot\vec{r}-\omega t\right)\right]$$

where the mode-labels and polarization vector are suppressed, and $\omega, \vec{k}$ are the angular frequency and wavevector of the traveling waves. The creation $\hat{b}^+$ and annihilation $\hat{b}$ operatators are related to the "position" $\hat{Q}$ and "momentum" $\hat{P}$ operators according to

$$\hat{Q} = \frac{1}{\sqrt{2}}\left[\hat{b} + \hat{b}^+\right] \qquad \hat{P} = \frac{-i}{\sqrt{2}}\left[\hat{b} - \hat{b}^+\right]$$

and they satisfy the commutation relations

$$\left[\hat{b},\hat{b}^+\right] = 1 \qquad \left[\hat{b},\hat{b}\right] = 0 = \left[\hat{b},\hat{b}^+\right]$$
$$\left[\hat{Q},\hat{P}\right] = i \qquad \left[\hat{Q},\hat{Q}\right] = 0 = \left[\hat{P},\hat{P}\right]$$

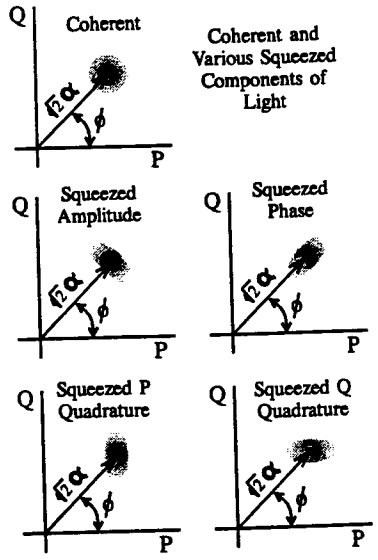


Figure 4.8.1: Various types of squeezed states as represented by the corresponding Wigner distribution. The coherent state has a value for the squeezing parameter of zero.

For coherent states $|\alpha\rangle$, the amplitude (to within a multiplicative constant) is defined by the complex parameter $\alpha = |\alpha| e^{i\phi}$ which can be related to the center of the Wigner probability distribution by

$$\alpha = \frac{1}{\sqrt{2}}\left(\bar{Q} + i\bar{P}\right)$$

where $\bar{Q} = \langle\alpha|\hat{Q}|\alpha\rangle = \sqrt{2}\,\mathrm{Re}(\alpha)$ and $\bar{P} = \langle\alpha|\hat{P}|\alpha\rangle = \sqrt{2}\,\mathrm{Im}(\alpha)$.

Squeezed electromagnetic states are characterized by reduced noise in one parameter but with added noise in the conjugate parameter. Figure 4.8.1 shows various types of squeezing as represented by a Wigner probability distribution plot. The amplitude-squeezed light, for example, has decreased amplitude variance and increased phase variance; the reverse is true for phase-squeezed light. For P-quadrature squeezed light, the possible range

of "P", denoted by ΔP, decreases while the range of "Q", denoted by ΔQ, increases; however, the product $\Delta Q\,\Delta P = 1/2$ remains unaltered. Making repeated measurements of the electric field for squeezed light would show that its amplitude and phase take on a range of values characterized by the ovals in the figure.

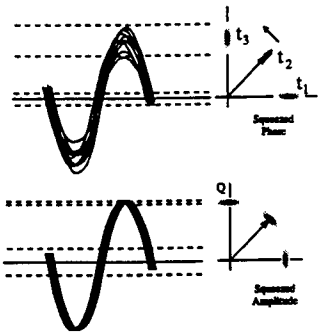


Figure 4.8.2: The top portion represents squeezed phase and the bottom represent squeezed amplitude.

The squeezed states can be viewed as sine waves by allowing the ovals to rotate and watching each point (Q,P) in the oval trace out a sine wave. Figure 4.8.2 shows traces for both phase and number squeezed light. Phase squeezed waves have almost the same phase but the amplitude can vary quite a bit (look at the crest of the sine wave). Number squeezed light has similar amplitudes but the phase can vary over a wide range (look at the nodes of the sine wave). The vacuum state can also be squeezed; the results appear similar to those in the right hand figure except that the average amplitude is zero. All squeezed electromagnetic waves are related to coherent states and, in particular, to squeezed vacuum states.

There are two sets of mathematical operations that produce squeezed states. Let $\hat{S}(\eta)$ denote an operator that "squeezes" a coherent state and let $\hat{D}(\alpha)$ be the displacement operator that defines a coherent state $|\alpha\rangle = \hat{D}(\alpha)|0\rangle$ from the vacuum. The parameter $\eta = re^{i\theta}$ gives the degree of squeezing where r>0. The first set of operations consists of squeezing the vacuum

$$\hat{S}(\eta)|0\rangle = |0,\eta\rangle \tag{4.8.1}$$

and then displacing the "new vacuum" through the phase-space distance α

$$\hat{D}(\alpha)\hat{S}(\eta)|0\rangle = |\alpha,\eta\rangle \tag{4.8.2}$$

The word "new vacuum" appears in quotes because it is not actually a physical vacuum; it is more similar to a two photon state. The squeezed states represented by Equation 4.8.2 are the type ploted in the above phase-space figures. The second set of operations consists of first displacing the vacuum and then squeezing the resulting coherent state (the reverse of Equation 4.8.2). This second type of squeezed state is not considered further because the final result does not have as simple an interpretation as the first type.

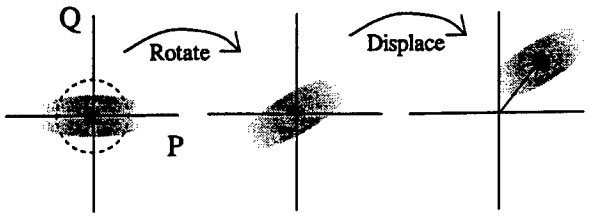


Figure 4.8.3: The vacuum is squeezed and then displaced to produce the squeezed coherent state.

The next sections provide the mathematical detail on squeezed states. The discussion starts with Q-squeezed vacuums using $\theta = 0$ in the squeeze parameter $\eta = re^{i\theta}$ so that "P" is the "long" axis. The length of the oval along the P-axis is a *factor* of $e^r/\sqrt{2}$ longer than the diameter of the circle representing the vacuum state $|0\rangle$; the Q-axis is "shortened" by the factor $e^{-r}/\sqrt{2}$. The Q-squeezed vacuum can be rotated to any

desired angle. The rotated, squeezed vacuum can be displaced to a new location by using the displacement operator $D(\alpha)$. The modulus of the parameter $\eta = r e^{i\theta}$ determines the amount of squeeze and the angle determines the axis of squeezing (i.e., amplitude or phase squeezed etc.).

Section 4.9: The Squeezing Operator and Squeezed States

The squeezing operator $\hat{S}(\eta)$ is the mathematical object that transforms a coherent state into a squeezed state. The complex parameter η determines the type of squeezed state which can be quadrature, number or phase squeezed. The probability distribution of the number-squeezed state (termed sub-poisson) is characterized by a standard deviation smaller than that for the coherent state but larger than for the Fock state. An anti-squeezed number state (phase squeezed) obeys super-poisson statistics. This section explores the squeezing operator $\hat{S}(\eta)$ along with some of its elementary properties.

The squeezing operator is defined by

$$\hat{S}(\eta) = \exp\left(\frac{\eta^*}{2}\hat{b}^2 - \frac{\eta}{2}\hat{b}^{+2}\right) \tag{4.9.1}$$

where $\eta = re^{i\theta}$. Unfortunately, it is not convenient to divide the squeezing operator into two or three terms using the Cambell-Baker-Hausdorff operator identity. However, it is possible to find useful relations that allows one to determine the variance of squeezed operators (such as electric field and energy), along with coordinate representations and photon statistics. As is evident, $\hat{S}$ is a unitary operator with $\hat{S}^+(\eta) = \hat{S}(-\eta) = \hat{S}^{-1}(\eta)$. The term $\hat{b}^{+2} + \hat{b}^2$ in the argument of the exponential of the squeezing operator is similar to the Hamiltonian for a parametric amplifier, which is one of the devices capable of producing squeezed light. A subsequent section shows that homodyne sensor systems can be used to detect and measure the amount of squeezing.

Either the quantum state $|\psi\rangle$ in the Hilbert space or the operators acting on the space can be squeezed. Applying the operator $\hat{S}(\eta)$ squeezes the states; it is usually applied to the vacuum .

$$|0,\eta\rangle = \hat{S}(\eta)|0\rangle$$

As an alternative to squeezed states, the operators can be squeezed by the similarity transformation $\hat{O}_S = \hat{S}^+\hat{O}\hat{S}$. It is the expectation value of an operator that is important.

$$\langle 0,\eta|\hat{O}|0,\eta\rangle = \langle 0|\left(\hat{S}^+\hat{O}\hat{S}\right)|0\rangle$$

For vacuum expectation values, either the states or the operators can be squeezed; however, one should not expect to find something "new" if both the operators and states are squeezed. The simplest method for evaluating an expectation value is to calculate the transformed operator $\hat{S}^+\hat{O}\hat{S}$. The operator $\hat{O}$ is assumed to be a functional of the creation and annihilation operators $\hat{O} = \hat{O}(\hat{b},\hat{b}^+)$. Therefore, the transformation of the operator $\hat{O}$ under $\hat{S}$ is known so long as the transformation of $\hat{b},\hat{b}^+$ is known. Then vacuum expectation value of the squeezed operator $\hat{O}_S$ can be evaluated.

What about squeezed coherent states? The simplest (but not the only) squeezed coherent state obtains by first squeezing the vacuum and then displacing the result.

$$|\alpha,\eta\rangle = \hat{D}(\alpha)\hat{S}(\eta)|0\rangle \tag{4.9.2}$$

where the coherent state without squeezing is $|\alpha\rangle = \hat{D}(\alpha)|0\rangle$. Note that the order of parameters in $|\alpha,\eta\rangle$ match the order of the operators in Equation 4.9.2.

Expectation values (such as average electric field or average energy) take the form

$$\langle\alpha,\eta|\hat{O}|\alpha,\eta\rangle = \langle 0|\hat{S}^{+}\hat{D}^{+}\hat{O}\hat{D}\hat{S}|0\rangle$$

Either the commutation relations for the displacement and squeezing operators must be calculated or the new operator $\hat{D}^{+}\hat{O}\hat{D}$ must be known. The product is particularly easy to calculate if $\hat{O} = \hat{O}\left(\hat{b},\hat{b}^{+}\right)$ since previous sections have shown

$$\hat{D}^{+}(\alpha)\hat{b}\hat{D}(\alpha) = \hat{b} + \alpha \qquad \text{and} \qquad \hat{D}^{+}(\alpha)\hat{b}^{+}\hat{D}(\alpha) = \hat{b}^{+} + \alpha^{*}$$

Products of the form $\hat{b}^{+}\hat{b}$ become

$$\hat{D}^{+}\left(\hat{b}^{+}\hat{b}\right)\hat{D} = \left(\hat{D}^{+}\hat{b}^{+}\hat{D}\right)\left(\hat{D}^{+}\hat{b}\hat{D}\right) = \left(\hat{b}^{+} + \alpha^{*}\right)\left(\hat{b} + \alpha\right)$$

(etc).

Eventually, one would like to calculate transition rates using squeezed coherent states as might be important for communication devices. For now, the discussion below calculates the variance of the electric field and electromagnetic Hamiltonian for the squeezed state.

Topic 4.9.1: The Squeezed Creation and Annihilation Operators

In order to find the transformation of more complicated operators, the first order of busines is to show

$$\hat{S}^{+}\hat{b}\hat{S} = \hat{b}\cosh(r) - \hat{b}^{+}e^{i\theta}\sinh(r) \qquad (4.9.3)$$
$$\hat{S}^{+}\hat{b}^{+}\hat{S} = \hat{b}^{+}\cosh(r) - \hat{b}e^{-i\theta}\sinh(r)$$

where $\eta = re^{i\theta}$ and $\hat{b}$, $\hat{b}^{+}$ are the annihilation and creation operators, respectively.

Consider the first of Equation 4.9.3. Expressions for the squeezed creation and annihilation operators are obtained by applying the operator expansion theorem from Chapter 2, which is

$$e^{-x\hat{A}}\hat{B}e^{x\hat{A}} = \hat{B} - \frac{x}{1!}[\hat{A},\hat{B}] + \frac{x^{2}}{2!}\left[\hat{A},[\hat{A},\hat{B}]\right] + \ldots$$

For the squeezing operator

$$\hat{S}(\eta) = \exp\left(\frac{\eta^{*}}{2}\hat{b}^{2} - \frac{\eta}{2}\hat{b}^{+2}\right)$$

the parameters are $\hat{A} = \frac{1}{2}\left(\eta^{*}\hat{b}^{2} - \eta\hat{b}^{+2}\right)$ and x=1. Substituting into the expansion formula provides

$$\hat{S}^{+}\hat{b}\hat{S} = e^{-\frac{1}{2}\left(\eta^{*}\hat{b}^{2} - \eta\hat{b}^{+2}\right)}\hat{b}e^{\frac{1}{2}\left(\eta^{*}\hat{b}^{2} - \eta\hat{b}^{+2}\right)}$$
$$= \hat{b} - \left[\frac{1}{2}\left(\eta^{*}\hat{b}^{2} - \eta\hat{b}^{+2}\right),\hat{b}\right] + \frac{1}{2!}\left[\frac{1}{2}\left(\eta^{*}\hat{b}^{2} - \eta\hat{b}^{+2}\right),\left[\frac{1}{2}\left(\eta^{*}\hat{b}^{2} - \eta\hat{b}^{+2}\right),\hat{b}\right]\right] + \ldots$$

The commutators in this expression are evaluated using $\left[\hat{b},\hat{b}^{+}\right] = 1$ to find

$$\hat{S}^{+}\hat{b}\hat{S} = \hat{b} - \eta^{*}\hat{b}^{+} + \frac{1}{2!}|\eta|^{2}\hat{b} - \frac{1}{3!}|\eta|^{2}\eta^{*}\hat{b}^{+} + \ldots$$

$$= \hat{b}\left(1 + \frac{1}{2!}r^{2} + \frac{1}{4!}r^{4} + \ldots\right) - \hat{b}^{+}e^{i\theta}\left(r + \frac{1}{3!}r^{3} + \ldots\right)$$

$$= \hat{b}\cosh(r) - \hat{b}^{+}e^{i\theta}\sinh(r)$$

which proves the first relation where "r" is the modulus of the squeezing parameter $re^{i\theta}$. Taking the adjoint proves the second relation.

Topic 4.9.2: The Squeezed EM Quadrature Operators

The electromagnetic "position" and "momentum" operators provide a first example of a squeezed operator that depends on the creation and annihilation operators. The "position and momentum" operators are defined by

$$\hat{Q} = \frac{1}{\sqrt{2}}\left(\hat{b} + \hat{b}^{+}\right) \qquad \hat{P} = \frac{1}{i\sqrt{2}}\left(\hat{b} - \hat{b}^{+}\right)$$

These operators prove to be important for plots of the Wigner probability distribution. Recall from the introduction to squeezed states (e.g., Figure 4.8.3) that the standard deviation of these operators is used to determine, in an easy way, the direction of squeezing.

Using Equation 4.9.3, the squeezed operator $\hat{Q}_S = \hat{S}^{+}\hat{Q}\hat{S}$ can be evaluated as

$$\hat{S}^{+}\hat{Q}\hat{S} = \frac{1}{\sqrt{2}}\hat{S}^{+}\left(\hat{b} + \hat{b}^{+}\right)\hat{S} = \frac{1}{\sqrt{2}}\left[\hat{S}^{+}\hat{b}\hat{S} + \hat{S}^{+}\hat{b}^{+}\hat{S}\right]$$

$$= \frac{1}{\sqrt{2}}\left[\hat{b}\cosh(r) - \hat{b}^{+}e^{i\theta}\sinh(r) + \hat{b}^{+}\cosh(r) - \hat{b}e^{-i\theta}\sinh(r)\right]$$

$$= \frac{\hat{b} + \hat{b}^{+}}{\sqrt{2}}\cosh(r) - \frac{\hat{b}e^{-i\theta} + \hat{b}^{+}e^{i\theta}}{\sqrt{2}}\sinh(r)$$

$$= \hat{Q}\cosh(r) - \hat{Q}_{R}\sinh(r)$$

The "rotated" operator $\hat{Q}_R$ is shorthand notation for

$$\hat{Q}_{R} = \frac{\hat{b}e^{-i\theta} + \hat{b}^{+}e^{i\theta}}{\sqrt{2}}$$

The rotated operator can also be written as

$$\hat{Q}_{R} = \hat{R}^{+}\hat{Q}\hat{R}$$

where $\hat{R} = e^{-i\hat{N}\theta} = e^{-i\hat{b}^{+}\hat{b}\theta}$ as discussed in Section 3.7.

Similarly, the squeezed momentum operator $\hat{P}_S = \hat{S}^{+}\hat{P}\hat{S}$ can be evaluated

$$\hat{S}^{+}\hat{P}\hat{S} = \frac{1}{i\sqrt{2}}\hat{S}^{+}\left(\hat{b} - \hat{b}^{+}\right)\hat{S} = \frac{\hat{b} - \hat{b}^{+}}{i\sqrt{2}}\cosh(r) - \frac{\hat{b}e^{-i\theta} - \hat{b}^{+}e^{i\theta}}{i\sqrt{2}}\sinh(r)$$

$$= \hat{P}\cosh(r) - \hat{P}_{R}\sinh(r)$$

Topic 4.9.3: Variance of the EM Quadrature

The operator for single mode electric fields can be written as

$$\vec{E}_k(\vec{r},t) = -\sqrt{\frac{\hbar\omega}{\varepsilon_o V}}\left[\hat{Q}\sin(\vec{k}\cdot\vec{r}-\omega t)+\hat{P}\cos(\vec{k}\cdot\vec{r}-\omega t)\right]$$

where the mode subscripts and polarization vector are suppressed, and $\omega, \vec{k}$ are the angular frequency and wavevector of the traveling waves. The "position" $\hat{Q}$ and "momentum" $\hat{P}$ operators are defined by

$$\hat{Q} = \frac{1}{\sqrt{2}}\left[\hat{b}+\hat{b}^+\right] \qquad \hat{P} = \frac{1}{i\sqrt{2}}\left[\hat{b}-\hat{b}^+\right]$$

and these operators satisfy the commutation relations

$$\left[\hat{Q},\hat{P}\right]=i \qquad \left[\hat{Q},\hat{Q}\right]=0=\left[\hat{P},\hat{P}\right] \qquad \left[\hat{b},\hat{b}^+\right]=1 \qquad \left[\hat{b},\hat{b}\right]=0=\left[\hat{b},\hat{b}^+\right]$$

The commutation relations imply that the two quadrature terms in the electric field cannot be simultaneously and precisely measured.

The coherent state for the electric field is normally obtained by displacing the vacuum; the resulting state is sometimes termed the "ideal" squeezed coherent state. The displacement operator does not change the "amount of squeezing". Therefore, one obtains a pictorial representation of the electromagnetic field in the squeezed vacuum by calculating the variance of the two quadratures. The Heisenberg uncertainty relation continues to hold regardless of the value of the squeezing parameter η.

First calculate the variance of $\hat{Q}$ for the squeezed vacuum state $|0,\eta\rangle = \hat{S}(\eta)|0\rangle$ where $\eta = re^{i\theta}$. The variance is

$$\sigma_Q^2 = \langle 0,\eta|\hat{Q}^2|0,\eta\rangle - \langle 0,\eta|\hat{Q}|0,\eta\rangle^2$$

The average of the "position" operator $\hat{Q}$ can be evaluated using Equations 4.9.3

$$\langle 0,\eta|\hat{Q}|0,\eta\rangle = \langle 0|\hat{S}^+\hat{Q}\hat{S}|0\rangle = \langle 0|\left\{\frac{\hat{b}+\hat{b}^+}{\sqrt{2}}\cosh(r)-\frac{\hat{b}e^{-i\theta}+\hat{b}^+e^{i\theta}}{\sqrt{2}}\sinh(r)\right\}|0\rangle = 0$$

Next, evaluate $\langle 0,\eta|\hat{Q}^2|0,\eta\rangle$ using $\hat{S}\hat{S}^+ = 1 = \hat{S}^+\hat{S}$

$$\begin{aligned}\langle 0,\eta|\hat{Q}^2|0,\eta\rangle &= \langle 0|\hat{S}^+\hat{Q}^2\hat{S}|0\rangle = \langle 0|\hat{S}^+\hat{Q}\hat{S}\,\hat{S}^+\hat{Q}\hat{S}|0\rangle \\ &= \langle 0|\left[\hat{Q}\cosh(r)-\hat{Q}_R\sinh(r)\right]\left[\hat{Q}\cosh(r)-\hat{Q}_R\sinh(r)\right]|0\rangle \\ &= \langle 0|\hat{Q}^2|0\rangle\cosh^2(r)+\langle 0|\hat{Q}_R^2|0\rangle\sinh^2(r)-\langle 0|\left(\hat{Q}\hat{Q}_R+\hat{Q}_R\hat{Q}\right)|0\rangle\cosh(r)\sinh(r)\end{aligned}$$

Substituting for $\hat{Q}$ and $\hat{Q}_R$ and using the commutation relations for the creation and annihilation operators we find the four relations

$$\langle 0|\hat{Q}^2|0\rangle = \langle 0|\left(\frac{\hat{b}+\hat{b}^+}{\sqrt{2}}\right)^2|0\rangle = \langle 0|\frac{\hat{b}^2+\hat{b}^{+2}+\hat{b}\hat{b}^++\hat{b}^+\hat{b}}{2}|0\rangle$$

$$= \frac{1}{2}\left\{\langle 0|\hat{b}^2|0\rangle+\langle 0|\hat{b}^{+2}|0\rangle+\langle 0|\hat{b}\hat{b}^+|0\rangle+\langle 0|\hat{b}^+\hat{b}|0\rangle\right\} = 0+0+\frac{1}{2}+0 = \frac{1}{2}$$

Similarly, one can show

$$\langle 0|\hat{Q}_R^2|0\rangle = \langle 0|\left(\frac{\hat{b}e^{-i\theta}+\hat{b}^+e^{i\theta}}{\sqrt{2}}\right)^2|0\rangle = \frac{1}{2}$$

$$\langle 0|\hat{Q}\hat{Q}_R|0\rangle = \langle 0|\left(\frac{\hat{b}+\hat{b}^+}{\sqrt{2}}\right)\left(\frac{\hat{b}e^{-i\theta}+\hat{b}^+e^{i\theta}}{\sqrt{2}}\right)|0\rangle = \frac{1}{2}\langle 0|\hat{b}\hat{b}^+e^{i\theta}+\hat{b}^+\hat{b}e^{-i\theta}|0\rangle = \frac{e^{i\theta}}{2}$$

$$\langle 0|\hat{Q}_R\hat{Q}|0\rangle = \frac{1}{2}\langle 0|\hat{b}^+\hat{b}e^{i\theta}+\hat{b}\hat{b}^+e^{-i\theta}|0\rangle = \frac{e^{-i\theta}}{2}$$

Substituting all of these into

$$\langle 0,\eta|\hat{Q}^2|0,\eta\rangle = \langle 0|\hat{Q}^2|0\rangle\cosh^2(r)+\langle 0|\hat{Q}_R^2|0\rangle\sinh^2(r)-\langle 0|\left(\hat{Q}\hat{Q}_R+\hat{Q}_R\hat{Q}\right)|0\rangle\cosh(r)\sinh(r)$$

provides

$$\langle 0,\eta|\hat{Q}^2|0,\eta\rangle = \frac{1}{2}\left\{\cosh^2(r)+\sinh^2(r)-2\cos(\theta)\sinh(r)\cosh(r)\right\}$$

Therefore the variance of $\hat{Q}$ is

$$\sigma_Q^2 = \langle 0,\eta|\hat{Q}^2|0,\eta\rangle - \langle 0,\eta|\hat{Q}|0,\eta\rangle^2 = \frac{1}{2}\left\{\cosh^2(r)+\sinh^2(r)-2\cos(\theta)\sinh(r)\cosh(r)\right\}$$

where $\eta = re^{i\theta}$. The variance of the "momentum" operator is similarly shown to be

$$\sigma_P^2 = \langle 0,\eta|\hat{P}^2|0,\eta\rangle = \frac{1}{2}\left\{\cosh^2(r)+\sinh^2(r)+2\cos(\theta)\sinh(r)\cosh(r)\right\}$$

Now consider the special case of $\theta = 0$. The variance of the EM quadratures operators become

$$\sigma_Q = \frac{e^{-r}}{\sqrt{2}} \quad \text{and} \quad \sigma_P = \frac{e^{r}}{\sqrt{2}}$$

Figure 4.9.1 shows the Wigner distribution for the squeezed vacuum. The greatest variance is along the "P" axis while the least is along the "Q" axis. Notice that the squeeze parameter "r" is always positive.

The angle θ in the squeezing parameter $\eta = re^{i\theta}$ controls the "direction" of squeezing. The "long" axis of squeezing is rotated by an angle $\theta/2$ as shown in the figure.

An important point is that the Heisenberg uncertainty relation remains unchanged for squeezed versus non-squeezed coherent states. For the squeezed vacuum, the Heisenberg uncertainty relation for the two quadrature components becomes

$$\sigma_Q\sigma_P = \frac{e^{-r}}{\sqrt{2}}\frac{e^{r}}{\sqrt{2}} = \frac{1}{2}$$

Even though the noise is squeezed out of one quadrature, it reappears in the other. However, it is now shown that the noise in the total electric field and Hamiltonian never decreases. Squeezing the vacuum always increases the variance of the Hamiltonian and total electric field (both quadratures).

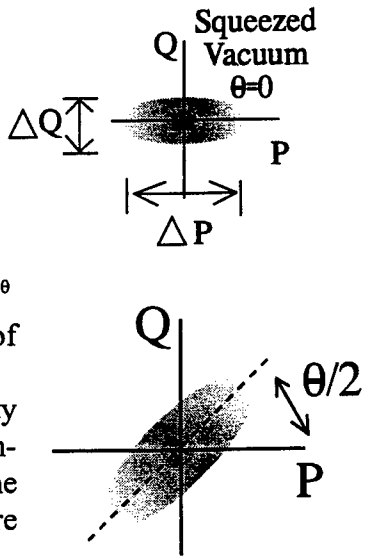


Figure 4.9.1: Top: The Q-quadrature is squeezed. Bottom: The angle θ in $\eta = re^{i\theta}$ rotates the "squeezed direction" by $\theta/2$.

Topic 4.9.4: Coordinate Representation of Squeezed States

The squeezed vacuum can be represented in either the "position" or "momentum" representation. Figure 4.9.2 shows the coordinate projection of the coherent state by the dotted lines. The solid lines represent the projection of the squeezed vacuum on to the coordinates. For example, the Gaussian distribution in the "P" coordinate is wider for the squeezed vacuum than it is for the coherent-state vacuum. The squeezed vacuum is seen to have a Gaussian distribution for either the "P" or the "Q" direction.

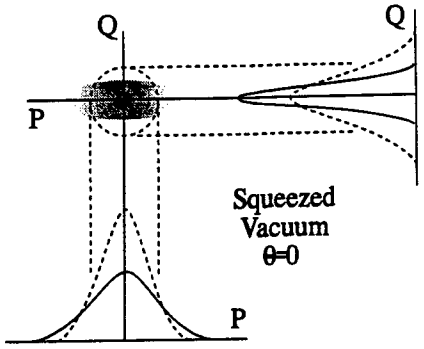


Figure 4.9.2: Comparison of a squeezed state (oval) and a coherent state (circle).

It is possible to show (Leonhardt) that the coordinate representation of the squeezed vacuum with the squeezing parameter $\eta = r e^{i\theta}$ (with $\theta = 0$) is given by

$$\Psi(Q) = e^{r/2}\, U_o\left(e^{r} Q\right) \tag{4.9.3}$$

and

$$\Psi'(P) = e^{-r/2}\, U_o'\left(e^{-r} P\right)$$

where $U_o(Q)$ is the coordinate representation of the coherent-state vacuum as given in Topic 4.5.3. The multiplicative factors $e^{\pm r/2}$ are for normalizing the wave functions. Leonhardt shows that Equation 4.9.3 leads to the correct form of the squeezing operator by differentiating with respect to "r" to get

$$\frac{\partial \Psi}{\partial r} = \frac{1}{2}\left(Q\frac{\partial}{\partial Q} + \frac{\partial}{\partial Q}Q\right)\psi = \frac{i}{2}\left(\hat{Q}\hat{P} + \hat{P}\hat{Q}\right)\Psi$$

This differential equation is solved by separating variables r, Ψ to find

$$\Psi(Q) = \exp\left\{\frac{ir}{2}\left(\hat{Q}\hat{P} + \hat{P}\hat{Q}\right)\right\}\Psi(0)$$

Therefore, Equation 4.9.2 is consistent with the squeezing operator for $\theta = 0$.

$$\hat{S} = \exp\left\{\frac{ir}{2}\left(\hat{Q}\hat{P} + \hat{P}\hat{Q}\right)\right\} = \exp\left\{\frac{r}{2}\left(\hat{b}^2 - \hat{b}^{+2}\right)\right\}$$

Section 4.10: Statistics for Squeezed States

As examples, this section calculates the average and variance of the electric field and the Hamiltonian for squeezed states. Also discussed is the probability distribution, which shows that photons occur in pairs for the squeezed state (sometimes the terms 2-photon state is applied). This section verifies further implicit claims made by the pictures in the introduction to squeezed states. One such claim is that the displacement operator does not affect the noise in the signal. Another is that the displaced vacuum state continues to be characterized by displacement vector $|\alpha\rangle$, despite the fact that the vacuum might be squeezed.

Topic 4.10.1: The Average Electric Field in a Squeezed Coherent State

The ideal squeezed coherent state is found by first squeezing the vacuum and then displacing the squeezed vacuum. As will be shown, the average of the electric field operator is unaffected by the squeezing (i.e., the average electric field in the ideal squeezed coherent state is identical to the average electric field in the coherent state). However, squeezing increases the variance of the *total* electric field more than that for the coherent state alone. As will be seen, the displacement operator does not affect the amount of noise in the field.

The quantized electric field operator is given by

$$\hat{E} = +i\sqrt{\frac{\hbar\omega}{2\varepsilon_o V}}\left[\hat{b}e^{ikx-i\omega t} - \hat{b}^+e^{-ikx+i\omega t}\right] \qquad (4.10.1)$$

The average electric field in the ideal squeezed coherent state $|\alpha,\eta\rangle = \hat{D}(\alpha)\hat{S}(\eta)|0\rangle$ is

$$\langle\alpha,\eta|\hat{E}|\alpha,\eta\rangle = \langle 0|\hat{S}^+\hat{D}^+\hat{E}\,\hat{D}\hat{S}|0\rangle$$

$$= +i\sqrt{\frac{\hbar\omega}{2\varepsilon_o V}}\left\{\langle 0|\hat{S}^+\left(\hat{D}^+\hat{b}\,\hat{D}\right)\hat{S}|0\rangle e^{ikx-i\omega t} - \langle 0|\hat{S}^+\left(\hat{D}^+\hat{b}^+\,\hat{D}\right)\hat{S}|0\rangle e^{-ikx+i\omega t}\right\}$$

The diplacement operator transforms the annihilation and creation operators according to

$$\hat{D}^+(\alpha)\hat{b}\,\hat{D}(\alpha) = \hat{b} + \alpha \qquad \text{and} \qquad \hat{D}^+(\alpha)\hat{b}^+\,\hat{D}(\alpha) = \hat{b}^+ + \alpha^*$$

so that

$$\langle\alpha,\eta|\hat{E}|\alpha,\eta\rangle = +i\sqrt{\frac{\hbar\omega}{2\varepsilon_o V}}\left\{\langle 0|\hat{S}^+\left(\hat{b}+\alpha\right)\hat{S}|0\rangle e^{ikx-i\omega t} - \langle 0|\hat{S}^+\left(\hat{b}^+ + \alpha^*\right)\hat{S}|0\rangle e^{-ikx+i\omega t}\right\}$$

Using the fact that the squeezing operator S is unitary, substituting for $\hat{S}^+\hat{b}\hat{S}$ (etc) and using $\langle 0|(b\cosh(r))|0\rangle = 0$ (etc.), we find

$$\langle\alpha,\eta|\hat{E}|\alpha,\eta\rangle = +i\sqrt{\frac{\hbar\omega}{2\varepsilon_o V}}\left\{\alpha\, e^{ikx-i\omega t} - \alpha^* e^{-ikx+i\omega t}\right\} = \langle\alpha|\hat{E}|\alpha\rangle$$

The average electric field is independent of the squeezing. Therefore the center of the Wigner distribution is independent of squeezing.

Topic 4.10.2: The Variance of the Electric Field in a Squeezed Coherent State

The variance of the *quadrature* components can be evaluated for the squeezed states in a manner similar to the average calculated for the electric field in Topic 4.10.1. The calculation would indicate that under the action of the displacement, the variance of the quadratures is invariant. That is

$$\langle\alpha,\eta|\left[\hat{P}^2-\overline{\overline{P}}^2\right]|\alpha,\eta\rangle=\langle 0,\eta|\left[\hat{P}^2-\overline{P}^2\right]|0,\eta\rangle$$

with a similar relation for $\hat{Q}$. The quantities $\overline{\overline{P}},\overline{P}$ are the average of $\hat{P}$ in the "displaced squeezed vacuum" and the "squeezed vacuum", respectively. The equation indicates that the degree of squeezing is independent of the displacement. Of course, to fully prove that the statistics of the coherent-state and the squeezed-coherent-states are independent of the average field amplitude α, it is only necessary to show that the character of the probability distribution is independent of the field amplitude (refer to Section 11 on the Wigner distribution).

This section calculates the variance of the total electric field, which also indicates that displacing the squeezed vacuum does not change the noise content. The calculation compares the total-field variances for the squeezed and coherent states at the end. For simplicity, use Equation 4.10.1 with x=0 and t=0 to calculate the variance of the electric field

$$\hat{E}\ =+i\sqrt{\frac{\hbar\omega}{2\varepsilon_o V}}\left[\hat{b}-\hat{b}^+\right]$$

The variance is given by

$$\sigma_E^2=\langle\alpha,\eta|\hat{E}^2|\alpha,\eta\rangle-\langle\alpha,\eta|\hat{E}|\alpha,\eta\rangle^2$$

The average is known from the previous page, which gives

$$\langle\alpha,\eta|\hat{E}|\alpha,\eta\rangle^2=\left\{i\sqrt{\frac{\hbar\omega}{2\varepsilon_o V}}\left[\alpha-\alpha^*\right]\right\}^2=-\frac{\hbar\omega}{2\varepsilon_o V}\left[\alpha-\alpha^*\right]^2 \qquad (4.10.2)$$

Next calculate the expectation of the square of the field

$$\langle\alpha,\eta|\hat{E}^2|\alpha,\eta\rangle=-\frac{\hbar\omega}{2\varepsilon_o V}\langle 0|\hat{S}^+\hat{D}^+\left(\hat{b}-\hat{b}^+\right)^2\hat{D}\hat{S}|0\rangle$$

$$=-\frac{\hbar\omega}{2\varepsilon_o V}\langle 0|\hat{S}^+\hat{D}^+\left(\hat{b}^2+\hat{b}^{+2}-\hat{b}\hat{b}^+-\hat{b}^+\hat{b}\right)\hat{D}\,\hat{S}|0\rangle$$

Inserting terms like

$$\hat{D}^+\hat{b}^2\hat{D}=\left(\hat{D}^+\hat{b}\hat{D}\right)\left(\hat{D}^+\hat{b}\hat{D}\right)=\left(\hat{b}+\alpha\right)\left(\hat{b}+\alpha\right) \quad \text{and} \quad \hat{D}^+\hat{b}\hat{b}^+\hat{D}=\left(\hat{b}+\alpha\right)\left(\hat{b}+\alpha^*\right) \quad \text{(etc)}$$

provides

$$\langle\alpha,\eta|\hat{E}^2|\alpha,\eta\rangle=-\frac{\hbar\omega}{2\varepsilon_o V}\langle 0|\hat{S}^+\left\{\left(\hat{b}+\alpha\right)^2+\left(\hat{b}^++\alpha^*\right)^2-\left(\hat{b}+\alpha\right)\left(\hat{b}^++\alpha^*\right)-\left(\hat{b}^++\alpha^*\right)\left(\hat{b}+\alpha\right)\right\}\hat{S}|0\rangle$$

Multiplying the terms and remembering to use the commutation relations, we find

$$\langle\alpha,\eta|\hat{E}^2|\alpha,\eta\rangle=-\frac{\hbar\omega}{2\varepsilon_o V}\langle 0|\hat{S}^+\left(\hat{b}-\hat{b}^+\right)^2\hat{S}|0\rangle-\frac{\hbar\omega}{2\varepsilon_o V}\left(\alpha-\alpha^*\right)^2 \qquad (4.10.3)$$

Evaluate the first term on the right hand side.

$$\langle 0|\hat{S}^{+}\left(\hat{b}-\hat{b}^{+}\right)^{2}\hat{S}|0\rangle=\langle 0|\hat{S}^{+}\left(\hat{b}^{2}+\hat{b}^{+2}-\hat{b}\hat{b}^{+}-\hat{b}^{+}\hat{b}\right)\hat{S}|0\rangle$$

Using the unitary property of "S" on terms such as $\hat{S}^{+}\hat{b}\hat{b}^{+}\hat{S}=\hat{S}^{+}\hat{b}\hat{S}\,\hat{S}^{+}\hat{b}^{+}\hat{S}=\hat{b}_{S}\hat{b}_{S}^{+}$, provides

$$\langle 0|\hat{S}^{+}\left(\hat{b}-\hat{b}^{+}\right)^{2}\hat{S}|0\rangle=\langle 0|\left\{\hat{b}_{S}^{2}+\hat{b}_{S}^{+2}-\hat{b}_{S}\,\hat{b}_{S}^{+}-\hat{b}_{S}^{+}\hat{b}_{S}\right\}|0\rangle \qquad (4.10.4)$$

Substituting for the squeezed annihilation and creation operators from Equations 4.9.3 which are

$$\hat{b}_{S}=\hat{S}^{+}\,\hat{b}\,\hat{S}=\hat{b}\cosh(r)-\hat{b}^{+}e^{i\theta}\sinh(r)$$

$$\hat{b}_{S}^{+}=\hat{S}^{+}\,\hat{b}^{+}\,\hat{S}=\hat{b}^{+}\cosh(r)-\hat{b}\,e^{-i\theta}\sinh(r)$$

Simplifying Equation 4.10.4 by noting "squared" terms like $\langle 0|\hat{b}^{2}|0\rangle\cosh(r)=0$ produce zero, as do terms with the destruction operator on the right $\langle 0|\,\hat{b}^{+}\hat{b}|0\rangle=0$, we find

$$\langle 0|\hat{S}^{+}\left(\hat{b}-\hat{b}^{+}\right)^{2}\hat{S}|0\rangle=\langle 0|\left\{-2\hat{b}\hat{b}^{+}\cos(\theta)-\hat{b}\hat{b}^{+}\cosh^{2}(r)+\hat{b}\hat{b}^{+}\sinh^{2}(r)\right\}|0\rangle$$

Recall the elementary relation $\cosh^{2}r-\sinh^{2}r=1$ and the fact that $\langle 0|\hat{b}\hat{b}^{+}|0\rangle=1$ to find

$$\langle 0|\hat{S}^{+}\left(\hat{b}-\hat{b}^{+}\right)^{2}\hat{S}|0\rangle=-2\cos(\theta)\cosh(r)\sinh(r)-1 \qquad (4.10.5)$$

Therefore, combining Equations 4.10.2 through 4.10.5, the variance of the electric field in the ideal coherent state becomes

$$\sigma_{E}^{2}=\langle\alpha,\eta|\hat{E}^{2}|\alpha,\eta\rangle-\langle\alpha,\eta|\hat{E}|\alpha,\eta\rangle^{2}=\frac{\hbar\omega}{2\varepsilon_{o}V}\left(2\cos(\theta)\cosh(r)\sinh(r)+1\right) \qquad (4.10.6)$$

where r>0 (and so sinh is always greater than zero). Which is independent of the displacement parameter α, as is necessary for the displacement operator not to influence the noise content.

The total-field variance in the ideal squeezed coherent state is always larger than the variance of the field in the pure coherent state. The variance of the pure coherent state is found by substituting r=0 in Equation 4.10.6

$$r=0\ \rightarrow\ \sigma_{E}^{2}=\frac{\hbar\omega}{2\varepsilon_{o}V}=\sigma_{E}^{2}\Big|_{\text{coherent}}$$

Therefore

$$\sigma_{E}^{2}\Big|_{sqz}\geq\sigma_{E}^{2}\Big|_{\text{coherent}}$$

As an important note, the main advantage to squeezed states is that for a system, one quadrature component might be the one to carry the signal (as for a parametric amplifier). The system is designed to ignore the noisy component.

Topic 4.10.3: The Average of the Hamiltonian in a Squeezed Coherent State

The average energy (i.e., expectation value of the Hamiltonian) in a squeezed state is always larger than the average energy in the corresponding coherent state. The energy of the squeezed state caries the energy of the corresponding coherent state plus the energy required to squeeze the vacuum. In particular, the squeezed vacuum has an

average energy larger than the true vacuum. Squeezed states consist of pairs of photons rather than single ones. The squeezed-state variance for the Hamiltonian is always larger than the coherent-state variance.

The single mode electromagnetic Hamiltonian is

$$\hat{H} = \hbar\omega\left(\hat{b}^{+}\hat{b} + \frac{1}{2}\right)$$

where ω is the angular frequency of the mode. The expectation value of the energy in the ideal squeezed coherent state is

$$\langle\alpha,\eta|\hat{H}|\alpha,\eta\rangle = \hbar\omega\langle\alpha,\eta|\left(\hat{b}^{+}\hat{b} + \frac{1}{2}\right)|\alpha,\eta\rangle = \hbar\omega\langle\alpha,\eta|\hat{b}^{+}\hat{b}|\alpha,\eta\rangle + \frac{1}{2}\hbar\omega$$

The first term on the right-hand side can be written as

$$\langle\alpha,\eta|\hat{b}^{+}\hat{b}|\alpha,\eta\rangle = \langle 0|\hat{S}^{+}\hat{D}^{+}\hat{b}^{+}\hat{b}\hat{D}\hat{S}|0\rangle = \langle 0|\hat{S}^{+}\left(\hat{D}^{+}\hat{b}^{+}\hat{D}\right)\left(\hat{D}^{+}\hat{b}\hat{D}\right)\hat{S}|0\rangle$$

which uses the fact that the displacement operator "D" is unitary. Substituting for the displaced creation and annihilation operators

$$\hat{D}^{+}(\alpha)\hat{b}\hat{D}(\alpha) = \hat{b} + \alpha \qquad \text{and} \qquad \hat{D}^{+}(\alpha)\hat{b}^{+}\hat{D}(\alpha) = \hat{b}^{+} + \alpha^{*}$$

the last expression becomes

$$\begin{aligned}\langle\alpha,\eta|\hat{b}^{+}\hat{b}|\alpha,\eta\rangle &= \langle 0|\hat{S}^{+}\hat{D}^{+}\hat{b}^{+}\hat{b}\hat{D}\hat{S}|0\rangle = \langle 0|\hat{S}^{+}\left(\hat{b}^{+} + \alpha^{*}\right)\left(\hat{b} + \alpha\right)\hat{S}|0\rangle \\ &= \langle 0|\hat{S}^{+}\left(\hat{b}^{+}\hat{b} + \alpha\hat{b}^{+} + \alpha^{*}\hat{b}^{+} + |\alpha|^{2}\right)\hat{S}|0\rangle \\ &= \langle 0|\hat{S}^{+}\hat{b}^{+}\hat{b}\,\hat{S}|0\rangle + \alpha\langle 0|\hat{S}^{+}\hat{b}^{+}\hat{S}|0\rangle + \alpha^{*}\langle 0|\hat{S}^{+}\hat{b}\hat{S}|0\rangle + |\alpha|^{2}\end{aligned}$$

Noting that the squeezed creation and annihilation operators

$$\begin{aligned}\hat{S}^{+}\hat{b}\hat{S} &= \hat{b}\cosh(r) - \hat{b}^{+}e^{i\theta}\sinh(r) \\ \hat{S}^{+}\hat{b}^{+}\hat{S} &= \hat{b}^{+}\cosh(r) - \hat{b}e^{-i\theta}\sinh(r)\end{aligned} \qquad (4.9.3)$$

(from Equations 4.9.3) are linear combinations of the creation and annihilation operators, it is apparent that the vacuum expectation values of the squeezed operators are zero. Therefore

$$\begin{aligned}\langle\alpha,\eta|\hat{b}^{+}\hat{b}|\alpha,\eta\rangle &= \langle 0|\hat{S}^{+}\hat{b}^{+}\hat{b}\,\hat{S}|0\rangle + |\alpha|^{2} \\ &= \langle 0|\left(\hat{S}^{+}\hat{b}^{+}\hat{S}\right)\left(\hat{S}^{+}\hat{b}\,\hat{S}\right)|0\rangle + |\alpha|^{2}\end{aligned}$$

Substituting Equations 4.9.3, the average becomes

$$\langle\alpha,\eta|\hat{b}^{+}\hat{b}|\alpha,\eta\rangle = \langle 0|\left(\hat{b}^{+}\cosh(r) - \hat{b}e^{-i\theta}\sinh(r)\right)\left(\hat{b}\cosh(r) - \hat{b}^{+}e^{i\theta}\sinh(r)\right)|0\rangle + |\alpha|^{2}$$

The terms $\langle 0|\hat{b}^{2}|0\rangle$, $\langle 0|\hat{b}^{+2}|0\rangle$ give zero as do the terms $\langle 0|\hat{b}^{+}\hat{b}|0\rangle$. We are left with

$$\langle\alpha,\eta|\hat{b}^{+}\hat{b}|\alpha,\eta\rangle = \langle 0|\hat{b}\hat{b}^{+}|0\rangle\sinh^{2}(r) + |\alpha|^{2} = \sinh^{2}(r) + |\alpha|^{2}$$

Therefore the expectation value for the Hamiltonian in the ideal squeezed coherent state is

$$\langle\alpha,\eta|\hat{H}|\alpha,\eta\rangle = \hbar\omega\langle\alpha,\eta|\hat{b}^{+}\hat{b}|\alpha,\eta\rangle + \frac{1}{2}\hbar\omega = \hbar\omega\left\{\sinh^{2}(r) + |\alpha|^{2} + \frac{1}{2}\right\} \qquad (4.10.7)$$

Equation 4.10.7 is an interesting result in itself. Notice that r=0 gives the average energy in the coherent state $|\alpha\rangle$. Squeezing the coherent state (i.e., r>0) causes the average energy to increase. Equation 4.10.6 consists of three terms. The last term is the

vacuum energy, the middle is the energy stored in the coherent state (similar to the square of the electric field) and the first term is the energy due to squeezing. Especially note that the average energy for the squeezed vacuum is larger than for just the vacuum alone.

$$\langle 0,\eta|\hat{H}|0,\eta\rangle = \hbar\omega\left\{\sinh^2(r)+\frac{1}{2}\right\}$$

As shown in the next topic, the squeezed vacuum appears as a multi-photon state.

The average energy in the coherent state $|\alpha\rangle$ is

$$\langle\alpha|\hat{H}|\alpha\rangle = \hbar\omega\left\{|\alpha|^2+\frac{1}{2}\right\} \tag{4.10.8}$$

The average energy in an ideal squeezed coherent state (Equation 4.10.7) can be written in terms of the average energy in a coherent state as

$$\langle\alpha,\eta|\hat{H}|\alpha,\eta\rangle = \langle\alpha|\hat{H}|\alpha\rangle + \hbar\omega\ \sinh^2(r)$$

so that

$$\langle\alpha,\eta|\hat{H}|\alpha,\eta\rangle \geq \langle\alpha|\hat{H}|\alpha\rangle$$

The energy is larger for squeezed states because the act of squeezing requires energy which is stored in the state. As will be seen, the squeezed vacuum is actually a multi-photon state. In fact, the photons occur in pairs and are never found with an "odd" number.

Topic 4.10.4: Photon Statistics for the Squeezed State

The squeezed vacuum is a multi-photon state which consists of a linear combination of Fock states with an even number of photons; that is, the photons occur in pairs. The probability of finding an odd number of photons is zero. This behavior is a result of the fact that the creation and annihilation operators are squared in the argument of the exponential for the squeezing operator

$$\hat{S}(\eta) = \exp\left[\frac{\eta^*}{2}\hat{b}^2 - \frac{\eta}{2}\hat{b}^{+2}\right]$$

For example, consider the case of a "Q-squeezed" vacuum (i.e., $\theta = 0$ in the squeezing parameter $\eta = re^{i\theta}$). Expanding the squeezing operator in powers of "r", provides

$$|0,r\rangle = \hat{S}(r)|0\rangle = \exp\left[\frac{r}{2}\left(\hat{b}^2-\hat{b}^{+2}\right)\right]|0\rangle$$

$$\cong\left[1+\left(\frac{r}{2}\right)\left(\hat{b}^2-\hat{b}^{+2}\right)+\frac{1}{2!}\left(\frac{r}{2}\right)^2\left(\hat{b}^2-\hat{b}^{+2}\right)^2+\frac{1}{3!}\left(\frac{r}{2}\right)^3\left(\hat{b}^2-\hat{b}^{+2}\right)^3+\ldots\right]|0\rangle$$

The terms with the creation and annihilation operators can be expanded to provide

$$|0,r\rangle = \hat{S}(r)|0\rangle = \left[1-\frac{1}{\sqrt{2}}\left(\frac{r}{2}\right)^2+\ldots\right]|0\rangle+\left[-\sqrt{2}\left(\frac{r}{2}\right)+\left(\frac{r}{2}\right)^3\left(\frac{4+4!}{6\sqrt{2}}\right)+\ldots\right]|2\rangle+\left[\left(\frac{r}{2}\right)^2\frac{\sqrt{4!}}{\sqrt{2}}+\ldots\right]|4\rangle+\ldots$$

For r=0, the series reduces to the vacuum state without squeezing. Notice that the squeezed vacuum is a sum of even Fock states. The probability for finding "n" photons in the vacuum during a measurement is

$$\text{prob}(n) = \left|\langle n|0,r\rangle\right|^2$$

The probability for odd "n" is always zero.

Leonhardt calculates the probability by using the coordinate representation of the squeezed vacuum. The coordinate wavefunction is given in Topic 4.9.4 as Equation 4.9.3

$$\Psi(Q) = e^{r/2}\, U_o\left(e^r Q\right)$$

where U_o is the unsqueezed vacuum wavefunction (a Gaussian). The probability amplitude is therefore

$$\langle n|\hat{S}(r)|0\rangle = \int_{-\infty}^{\infty} dQ\, \psi_n(Q)\, e^{r/2}\, U_o\left(e^r Q\right)$$

where $\psi_n(Q)$ is the coordinate representation of the Fock wavefunction for a state consisting of "n" photons. Leonhardt gives the formula

$$\text{Prob}(n) = \begin{cases} 0 & n = 1,3,5\ldots \\ \binom{n}{n/2} \dfrac{1}{\cosh(r)} \left(\dfrac{\tanh(r)}{2}\right)^n & n = 0,2,4\ldots \end{cases}$$

where

$$\binom{n}{m} = \frac{n!}{m!\,(n-m)!}$$

The figure shows a plot of the probability of "n" photons for r=0.25 and r=1.

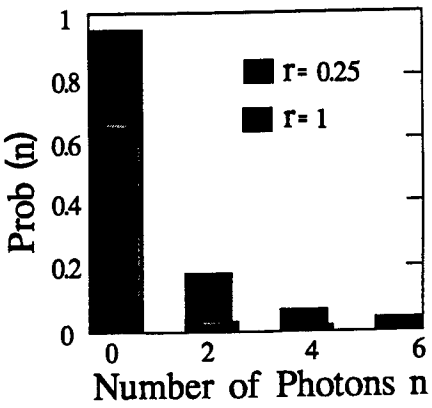


Figure 4.10.1: Probability distribution for finding "n" photons in the squeezed vacuum for two values of the squeezing parameter "r". A value of r=0.25 produces an almost unsqueezed state.

Section 4.11: The Wigner Distribution

Previous sections discuss the Fock, coherent and squeezed states of the electromagnetic field. The Fock state is a number-squeezed coherent state, which means that the probability distribution for the number of photons is "sub-Poisson" (i.e., the standard deviation for the number operator is smaller than that for the Poisson distribution). The quadrature components of the electric field cannot simultaneously and precisely be measured. A Gaussian distribution characterizes the coherent state in the position or momentum coordinate representation. Introductory material in previous sections suggests that a distribution in phase space can be defined that treats the phase space coordinates (position and momentum) as random variables. The Wigner distribution is the closest quantum analogue to the classical probability distribution. Unfortunately, the Wigner distribution can become negative (a non-classical probability property) for certain types of states (such as Fock states). For coherent and squeezed states, the Wigner distribution provides a near-classical picture. The so-called "P" and "Q" quasi-probability distributions can also be defined; these distributions are *not* related to the "P and Q coordinates". For the "P" distribution, a mathematical object very similar to a density operator is defined for coherent states.

Topic 4.11.1: The Wigner Formula and an Example

The Wigner distribution is defined to be the joint quasi-probability density for the EM field "position Q" and "momentum P" coordinates. This section shows that the Wigner function is given by

$$W(Q,P)=\frac{1}{2\pi}\int_{-\infty}^{\infty}dx\,\exp(iPx)\left\langle Q-\frac{x}{2}\right|\hat{\rho}\left|Q+\frac{x}{2}\right\rangle$$

where $\hat{\rho}$ is a density operator (see below). We can define

$$a=\frac{1}{\sqrt{2}}(Q+iP)$$

to be the (normalized) electric field amplitude which would be the annihilation operator if Q,P were operators. Notice that for the coherent state $|\alpha\rangle$, the value "a" is not the same as α. This is due to the fact that $\alpha=\frac{1}{\sqrt{2}}(\overline{Q}+\overline{P})$ and $\overline{Q},\overline{P}$ are the average values of $\hat{Q},\hat{P}$ in the coherent state whereas Q,P are measured values of $\hat{Q},\hat{P}$. Each set of measured values Q,P leads to a different sine wave since Q,P are related to an amplitude and phase for the EM field (see, for example, the Figures 4.3.6 and 4.8.2).

Example: Find the Wigner function for the vacuum state

The density operator represents any quantum mechanical state for which we wish to find the quasi-probability density. To fix our thoughts, find the Wigner function for the coherent state $|\alpha\rangle$ by setting $\hat{\rho}=|\alpha\rangle\langle\alpha|$ (the Wigner function for non-pure states can also be found). A term such as $\langle Q|\alpha\rangle$ in

$$\langle Q|\hat{\rho}|Q\rangle = \langle Q|\alpha\rangle\langle\alpha|Q\rangle = |\langle Q|\alpha\rangle|^2$$

is the coordinate representation of the wave function. Recall the coordinate wavefunction for the vacuum from Equations 4.5.13 and 4.5.14.

$$\langle Q|\alpha = 0\rangle = \psi_o(Q) = \frac{1}{\pi^{1/4}} e^{-Q^2/2}$$

The Wigner function is then

$$W(Q,P) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\, \exp(iPx)\left\langle Q - \frac{x}{2}\middle|\hat{\rho}\middle|Q + \frac{x}{2}\right\rangle$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\, \exp(iPx)\left\langle Q - \frac{x}{2}\middle|\alpha = 0\right\rangle\left\langle\alpha = 0\middle|Q + \frac{x}{2}\right\rangle$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\, \exp(iPx)\psi_o^*\left(Q + \frac{x}{2}\right)\psi_o\left(Q - \frac{x}{2}\right)$$

or, substituting for the wavefunctions

$$W(Q,P) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\, \exp(iPx)\frac{1}{\pi^{1/4}}\exp\left[-\frac{1}{2}\left(Q + \frac{x}{2}\right)^2\right]\frac{1}{\pi^{1/4}}\exp\left[-\frac{1}{2}\left(Q - \frac{x}{2}\right)^2\right]$$

$$= \frac{1}{\pi}\exp\left(-Q^2 - P^2\right)$$

The last line follows after some simplifying algebra and an integration. The Wigner function is a Gaussian centered on the phase-space origin as discussed in the introductory material in Sections 4.3 and 4.8.

The Wigner function connects the quantum theory and classical probability theory through the use of marginal probabilities. For a classical probability density function $W(x,y)$, the probability of finding (for example) the values of $x \in (a,b)$ and $y \in (c,d)$ are given by

$$Prob(a < x < b, c < y < d) = \int_a^b\int_c^d dx\, dy\, W(x,y)$$

The probability that $x \in (a,b)$ regardless of the value of "y" is

$$Prob(a < x < b) = Prob(a < x < b, -\infty < y < \infty) = \int_a^b\int_{-\infty}^{\infty} dx\, dy\, W(x,y)$$

Therefore, the probability density function for "x" is identified as

$$W(x) \equiv \int_{-\infty}^{\infty} dy\, W(x,y)$$

The Wigner formula is defined in such a way that simultaneous measurements of $\hat{P},\hat{Q}$ do not need to be considered. The Wigner formulation requires the density functions for Q and P be given by

$$W_Q(Q) \equiv \int_{-\infty}^{\infty} dP\, W(Q,P) = \langle Q|\hat{\rho}|Q\rangle \quad \text{and} \quad W_P(P) \equiv \int_{-\infty}^{\infty} dQ\, W(Q,P) = \langle P|\hat{\rho}|P\rangle$$

where the subscripts P and Q remind the reader of the independent variable which is not always obvious in what follows. The 1-D probability density functions are related to the density operator as

$$W_Q(Q) = \langle Q|\hat{\rho}|Q\rangle \sim \langle Q|\psi\rangle\langle\psi|Q\rangle = |\psi(Q)|^2$$

and similarly for the density that depends on the "P" coordinate. This formula is reminiscent of the "shadow" plots such as Figure 4.5.1.

Topic 4.11.2: Derivation of the Wigner Formula

The derivation follows that in Leonhardt's wonderful book with some notational changes. The argument uses the notation Q', P' for the rotated coordinates which reduce to Q and P for a rotation of 0 degrees (it's ok to think of Q,P instead however). The following list provides the sequence of steps required for the derivation. The function W' is the Wigner function of Q', P'; it reduces to W when the coordinates are not rotated.

1. The classical marginal probability distribution is defined as

$$W_Q(Q') \equiv \int_{-\infty}^{\infty} dP'\, W'(Q',P')$$

which has the form of a "Radon" transformation. The function $W'(Q',P')$ is the quantity of interest; i.e., it is the Wigner function. The inverse transformation of the previous equation must be developed in order to isolate $W'(Q',P')$.

2. The "characteristic function" is the Fourier Transform of the probability density $W'(Q',P')$ which appears in the integrand of Step 1. The Fourier transform will be found to be

$$\tilde{W}_Q(\xi,\theta) = \sqrt{2\pi}\,\tilde{W}(\xi\cos\theta, \xi\sin\theta) = \sqrt{2\pi}\,\tilde{W}(u,v)$$

where "~" denotes Fourier transform, ξ is the Fourier transform variable conjugate to Q', θ is the rotation angle for the coordinates and $u = \xi\cos\theta$ and $v = \xi\sin\theta$. The formula above relates the Fourier transform of the marginal probability $\tilde{W}_Q$ to the Fourier transform of Wigner function $\tilde{W}$. Essentially, this step provides the desired inversion for the Radon transfomation. The above characteristic function is to be compared with another expression obtained from quantum theory.

3. The quantum probability distribution is define through the density operator $\hat{\rho}$

$$W(Q') = \langle Q'|\hat{\rho}|Q'\rangle$$

and similar to Step 2, the Fourier transform provides

$$\tilde{W}_Q(\xi,\theta)=\frac{1}{\sqrt{2\pi}}\mathrm{Tr}\left\{\hat{\rho}\,\exp\left(-iu\hat{Q}-iv\hat{P}\right)\right\}$$

where the Weyl operator $\exp\left(-iu\hat{Q}-iv\hat{P}\right)$ is structurally similar to the displacement operator.

4. Equating the results of Steps 2 and 3 provides an expression for the Fourier transform of the Wigner distribution

$$\tilde{W}(u,v)=\frac{1}{2\pi}\mathrm{Tr}\left\{\hat{\rho}\,\exp\left(-iu\hat{Q}-iv\hat{P}\right)\right\}$$

The Wigner distribution is obtained by an inverse Fourier transform and using the wavefunction translation properties inherent to the Weyl operator.

The rest of the discussion executes the program plan in steps 1 through 4 above. First, as step 0, the rotated coordinates are discussed.

Step 0: The Rotations

The coordinates Q',P' are defined by the rotation

$$\begin{pmatrix}Q'\\P'\end{pmatrix}=\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}Q\\P\end{pmatrix}$$

Recall that the coordinate transformation is defined in such a way that a function of the new coordinates and a function of the old coordinates (refer to Appendix 5) are related by

$$W'(Q',P')=W(Q,P)$$

The W and W' are the Wigner functions in the original and rotated coordinates. The coordinates Q and P are the eigenvalues of the operators $\hat{Q},\hat{P}$ (respectively) according to

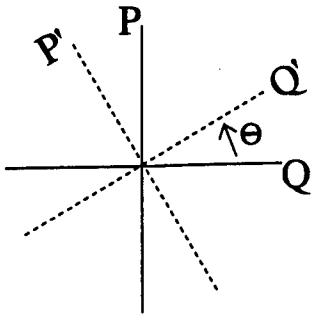


Figure 4.11.1: Definition of the rotated coordinates.

$$\hat{Q}|Q\rangle=Q|Q\rangle \qquad \text{and} \qquad \hat{P}|P\rangle=P|P\rangle$$

The rotation operator $\hat{R}$, which rotates the operators $\hat{Q},\hat{P}$, is defined by

$$\hat{R}=e^{-i\hat{N}\theta}$$

where $\hat{N}$ is the number operator $\hat{N}=\hat{b}^+\hat{b}$, and $\hat{b}^+,\hat{b}$ are the creation and annihilation operators respectively. The rotated "position" and "momentum" operators are

$$\hat{Q}'=\hat{R}^+\hat{Q}\hat{R} \qquad \text{and} \qquad \hat{P}'=\hat{R}^+\hat{P}\hat{R}$$

For example, $\hat{Q}'$ is found to be

$$\hat{Q}'=\hat{R}^+\hat{Q}\hat{R}=e^{i\hat{N}\theta}\frac{1}{\sqrt{2}}\left(\hat{b}+\hat{b}^+\right)e^{-i\hat{N}\theta}=\frac{1}{\sqrt{2}}\left(\hat{b}e^{-i\theta}+\hat{b}^+e^{i\theta}\right)=\hat{Q}\cos\theta+\hat{P}\sin\theta$$

Let the states be defined by

$$|Q,\theta\rangle\equiv|Q'\rangle=\hat{R}^+|Q\rangle \qquad \text{and} \qquad |P,\theta\rangle\equiv|P'\rangle=\hat{R}^+|P\rangle$$

(note the "$\hat{R}^+$"). The adjoint rotation operator $\hat{R}^+$, which is unitary, has the property that $\hat{R}^+(\theta) = \hat{R}(-\theta)$. The coordinate kets rotated in this manner are eigenkets of the rotated operator

$$\hat{Q}'|Q'\rangle = \left(\hat{R}^+\hat{Q}\hat{R}\right)\hat{R}^+|Q\rangle = \hat{R}^+\hat{Q}|Q\rangle = \hat{R}^+Q|Q\rangle = Q\hat{R}^+|Q\rangle = Q|Q'\rangle$$

Step 1: The Marginal Probability Distribution

The classical marginal probability distribution is defined as

$$W_Q(Q') \equiv \int_{-\infty}^{\infty} dP'\, W'(Q',P')$$

To find the "inverse" of the transformation, one needs to work with the Fourier transformation.

Step 2: The Characteristic Function

The Fourier transform of

$$W_Q(Q') \equiv \int_{-\infty}^{\infty} dP'\, W'(Q',P')$$

is given by

$$\tilde{W}_Q(\xi,\theta) = \int_{-\infty}^{\infty} dQ'\, W_Q(Q')\frac{e^{-i\xi Q'}}{\sqrt{2\pi}} = \int_{-\infty}^{\infty} dQ'\, \frac{e^{-i\xi Q'}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dP'\, W'(Q',P') = \iint_{R^2} dQ'dP'\, \frac{e^{-i\xi Q'}}{\sqrt{2\pi}} W'(Q',P')$$

The Jacobian of the transformation provides

$$dQ\, dP = dQ'\, dP'$$

and the definition for the rotation of a function (Appendix 5) gives

$$W'(Q',P') = W(Q,P)$$

and the coordinate rotation yields

$$Q' = Q\cos\theta + P\sin\theta$$

Therefore $\tilde{W}_Q(\xi,\theta)$ becomes

$$\begin{aligned}\tilde{W}_Q(\xi,\theta) &= \iint dQ\, dP\, \frac{e^{-i\xi(Q\cos\theta + P\sin\theta)}}{\sqrt{2\pi}} W(Q,P) \\ &= \sqrt{2\pi}\iint dQ\, dP\, W(Q,P)\frac{e^{-iQ(\xi\cos\theta) - iP(\xi\sin\theta)}}{2\pi}\end{aligned} \qquad (4.11.1)$$

The characteristic function is the Fourier transform of the probability density

$$\tilde{W}(u,v) = \iint dQ\, dP\, W(Q,P)\frac{e^{-iQu - iPv}}{2\pi}$$

where the transform variables are related to polar coordinates by

$$u = \xi\cos\theta \qquad \text{and} \qquad v = \xi\sin\theta$$

Therefore, Equation 4.11.1 relates the characteristic function to the Fourier transform of the marginal probability

$$\tilde{W}_Q(\xi,\theta) = \sqrt{2\pi}\,\tilde{W}(\xi\cos\theta,\xi\sin\theta) = \sqrt{2\pi}\,\tilde{W}(u,v)$$

Step 3: Fourier Transform of the Quantum Probability Density Function

The quantum probability density (in the Q' coordinate) is given by

$$W_Q(Q') = \langle Q'|\hat{\rho}|Q'\rangle = \langle Q|\hat{R}\hat{\rho}\hat{R}^+|Q\rangle$$

The Fourier transform of this function is

$$\tilde{W}_Q(\xi,\theta) = \int_{-\infty}^{\infty} dQ\,\langle Q|\hat{R}\hat{\rho}\hat{R}^+|Q\rangle\frac{e^{-i\xi Q}}{\sqrt{2\pi}} = \int_{-\infty}^{\infty} dQ\,\langle Q|\hat{R}\hat{\rho}\hat{R}^+\frac{e^{-i\xi Q}}{\sqrt{2\pi}}|Q\rangle$$

where the constant exponential term is moved inside the coordinate expectation value. Using $e^{-i\xi\hat{Q}}|Q\rangle = e^{-i\xi Q}|Q\rangle$ (as can be seen by Taylor expanding $e^{-i\xi\hat{Q}}$), we obtain

$$\tilde{W}_Q(\xi,\theta) = \int_{-\infty}^{\infty} dQ\,\langle Q|\hat{R}\hat{\rho}\hat{R}^+\frac{e^{-i\xi\hat{Q}}}{\sqrt{2\pi}}|Q\rangle = \frac{1}{\sqrt{2\pi}}\mathrm{Tr}\left[\hat{R}\hat{\rho}\hat{R}^+\frac{e^{-i\xi\hat{Q}}}{\sqrt{2\pi}}\right]$$

where the coordinate basis is used for the trace (see Topic 2.3.5). The order of the operators in the trace can be permuted (see Section 2.3) to get

$$\tilde{W}_Q(\xi,\theta) = \int_{-\infty}^{\infty} dQ\,\langle Q|\hat{R}\hat{\rho}\hat{R}^+\frac{e^{-i\xi\hat{Q}}}{\sqrt{2\pi}}|Q\rangle = \frac{1}{\sqrt{2\pi}}\mathrm{Tr}\left[\hat{\rho}\hat{R}^+e^{-i\xi\hat{Q}}\hat{R}\right] \qquad (4.11.2)$$

However, the previous discussion shows that $\hat{Q}' = \hat{R}^+\hat{Q}\hat{R} = \hat{Q}\cos\theta + \hat{P}\sin\theta$ and therefore, expanding the exponential, and using the property that $\hat{R}$ is unitary,

$$\hat{R}^+e^{-i\xi\hat{Q}}\hat{R} = \hat{R}^+\left(\sum_n\frac{1}{n!}(-i\xi)^n\hat{Q}^n\right)\hat{R} = \sum_n\frac{1}{n!}(-i\xi)^n\left(\hat{R}^+\hat{Q}\hat{R}\right)^n = e^{-i\xi\left(\hat{Q}\cos\theta+\hat{P}\sin\theta\right)}$$

Substituting this into Equation 4.11.2 along with

$$u = \xi\cos\theta \qquad \text{and} \qquad v = \xi\sin\theta$$

provides

$$\tilde{W}_Q(\xi,\theta) = \frac{1}{\sqrt{2\pi}}\mathrm{Tr}\left[\hat{\rho}\hat{R}^+e^{-i\xi\hat{Q}}\hat{R}\right] = \frac{1}{\sqrt{2\pi}}\mathrm{Tr}\left[\hat{\rho}\,e^{-i\hat{Q}u-i\hat{P}v}\right]$$

Step 4: Equate the results of Steps 2 and 3

Equating the results of Steps 2 and 3 provides

$$\tilde{W}(u,v) = \frac{1}{2\pi}\mathrm{Tr}\left[\hat{\rho}\,e^{-i\hat{Q}u-i\hat{P}v}\right] \qquad (4.11.3)$$

The Weyl operator can be rewritten using the Baker-Hausdorff formula

$$e^{\hat{A}+\hat{B}} = e^{\hat{A}}e^{\hat{B}}e^{-[\hat{A},\hat{B}]/2} \quad \text{so long as} \quad \left[\hat{A},\left[\hat{A},\hat{B}\right]\right] = 0 = \left[\hat{B},\left[\hat{A},\hat{B}\right]\right]$$

from Chapter 2. The Weyl operator becomes

$$e^{-i\hat{Q}u-i\hat{P}v} = e^{-i\hat{Q}u}e^{-i\hat{P}v}e^{-\left[-i\hat{Q}u,-i\hat{P}v\right]/2} = e^{-i\hat{Q}u}e^{-i\hat{P}v}e^{+iuv/2}$$

since $\left[\hat{Q},\hat{P}\right] = i$. Equation 4.11.3 can be written

$$\tilde{W}(u,v) = \frac{1}{2\pi} Tr\left[\hat{\rho}\, e^{-i\hat{Q}u - i\hat{P}v}\right] = \frac{1}{2\pi}\int_{-\infty}^{\infty} dQ \langle Q|\hat{\rho}\, e^{-iu\hat{Q} - iv\hat{P}}|Q\rangle$$
$$= \frac{e^{+iuv/2}}{2\pi}\int_{-\infty}^{\infty} dQ \langle Q|\hat{\rho} e^{-iu\hat{Q}} e^{-iv\hat{P}}|Q\rangle \qquad (4.11.4)$$

Using the fact that $e^{-iv\hat{P}}$ is a translation operator in Q space according to Section 2.16, provides

$$e^{-iv\hat{P}}|Q\rangle = |Q+v\rangle$$

so that

$$e^{-iu\hat{Q}} e^{-iv\hat{P}}|Q\rangle = e^{-iu\hat{Q}}|Q+v\rangle = e^{-iu(Q+v)}|Q+v\rangle$$

Equation 4.11.4 becomes

$$\tilde{W}(u,v) = \frac{e^{iuv/2}}{2\pi}\int_{-\infty}^{\infty} dQ \langle Q|\hat{\rho} e^{-iu(Q+v)}|Q+v\rangle = \frac{e^{-iuv/2}}{2\pi}\int_{-\infty}^{\infty} dQ\, e^{-iuQ} \langle Q|\hat{\rho}|Q+v\rangle$$

Making the substitution

$$x = Q - \frac{v}{2}$$

yields

$$\tilde{W}(u,v) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\, e^{-iux} \left\langle x - \frac{v}{2}\middle|\hat{\rho}\middle|x + \frac{v}{2}\right\rangle \qquad (4.11.5)$$

Step 5: Inverse Transform

The inverse transform of Equation 4.11.5 is

$$W(Q,P) = \iint_{(-\infty,\infty)} du\, dv\, \tilde{W}(u,v) \frac{e^{iuQ+ivP}}{\left(\sqrt{2\pi}\right)^2} = \frac{1}{2\pi}\iint dx\, dv \left\langle x - \frac{v}{2}\middle|\hat{\rho}\middle|x + \frac{v}{2}\right\rangle e^{ivP} \int_{-\infty}^{\infty} du\, \frac{e^{iu(Q-x)}}{2\pi}$$

The definition for the Dirac delta function

$$\int_{-\infty}^{\infty} du\, \frac{e^{iu(Q-x)}}{2\pi} = \delta(Q-x)$$

gives

$$W(Q,P) = \frac{1}{2\pi}\iint dx\, dv \left\langle x - \frac{v}{2}\middle|\hat{\rho}\middle|x + \frac{v}{2}\right\rangle e^{ivP}\, \delta(Q-x)$$

Finally, we obtain Wigner's formula for a quasi-probability distribution.

$$W(Q,P) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dv\, \exp(ivP) \left\langle Q - \frac{v}{2}\middle|\hat{\rho}\middle|Q + \frac{v}{2}\right\rangle$$

where Q and P are coordinates instead of operators.

Topic 4.11.3: Examples for the Wigner Function

Wigner's formula

$$W(Q,P)=\frac{1}{2\pi}\int_{-\infty}^{\infty}dx\,\exp(ixP)\left\langle Q-\frac{x}{2}\middle|\hat{\rho}\middle|Q+\frac{x}{2}\right\rangle$$

is easy to evaluate so long as the wave function $|\psi\rangle$ in $\hat{\rho}=|\psi\rangle\langle\psi|$ is known as a function of the "position" coordinate Q. The inner product then produces

$$\left\langle Q-\frac{x}{2}\middle|\hat{\rho}\middle|Q+\frac{x}{2}\right\rangle=\left\langle Q-\frac{x}{2}\middle|\psi\right\rangle\left\langle\psi\middle|Q+\frac{x}{2}\right\rangle=\psi\left(Q-\frac{x}{2}\right)\psi^{*}\left(Q+\frac{x}{2}\right)$$

Example 1: Wigner representation of the Fock state $|1\rangle$.

The coordinate representation of the first Fock state is

$$\langle Q|1\rangle=\psi_1(Q)=C_1 Q\exp\left(-\frac{Q^2}{2}\right)$$

from Equation 4.2.5. Therefore, set $\rho_1=|1\rangle\langle 1|$ to get

$$\left\langle Q-\frac{x}{2}\middle|\hat{\rho}\middle|Q+\frac{x}{2}\right\rangle=\left\langle Q-\frac{x}{2}\middle|1\right\rangle\left\langle 1\middle|Q+\frac{x}{2}\right\rangle=\psi_1\left(Q-\frac{x}{2}\right)\psi_1^{*}\left(Q+\frac{x}{2}\right)$$

which gives

$$\left\langle Q-\frac{x}{2}\middle|\hat{\rho}\middle|Q+\frac{x}{2}\right\rangle=|C_1|^2\left(Q-\frac{x}{2}\right)\exp\left(-\frac{(Q-x/2)^2}{2}\right)\left(Q+\frac{x}{2}\right)\exp\left(-\frac{(Q+x/2)^2}{2}\right)$$

The Wigner function becomes

$$W(Q,P)=\frac{|C_1|^2}{2\pi}\int_{-\infty}^{\infty}dx\,\exp(ixP)\left(Q-\frac{x}{2}\right)\exp\left(-\frac{(Q-x/2)^2}{2}\right)\left(Q+\frac{x}{2}\right)\exp\left(-\frac{(Q+x/2)^2}{2}\right)$$

which can be integrated.

Section 4.12: The "P" Quasi-Probability Distribution

The "P" distribution function is defined as P_α in

$$\hat{\rho}=\int d^2\alpha\, P_\alpha|\alpha\rangle\langle\alpha|$$

where P_α is clearly intended to take the place of a classical probability in the definition of the density operator. In the above version of the formula, $|\alpha\rangle$ is a coherent state, but the formula can be written for other wavefunctions.

Section 4.13: Measuring the Noise in Squeezed States

Homodyne detection is the most common method for measuring the electromagnetic noise in a squeezed state. There exists both balanced and unbalanced detectors which, for RF date back to WWII. Any detection system can, of course, detect the average number of photons in the beam. This section discusses a simple homodyne detection system.

Figure 4.13.1 shows a block diagram for producing and detecting optical electromagnetic waves in the squeezed state. A single photodetector could be used but it would detect the total light emission and not just a single quadrature component. A laser generates a large amplitude optical signal (which can be represented by a classical wave). A system, denoted by "Sqz", converts the laser signal to squeezed light. The squeezed light is used in any desired manner after it leaves the "squeezer"; the figure shows the squeezed light incident on a Device Under Test (DUT). A reference signal is derived through a beam splitter; this signal maintains a fixed phase relation with the squeezed light. The homodyne detection system consists of perfect 50-50 beam splitter (no surface reflections and no absorption), two photodetectors and a summer with one input inverted. By adjusting the phase θ of the reference signal, it is possible to measure the quadrature component $\hat{Q}_\theta = \hat{R}^+ \hat{Q} \hat{R}$ where the rotation operator is $\hat{R}(\theta) = e^{-i\hat{N}\theta}$ and $\hat{N} = \hat{b}^+\hat{b}$. Recall that $\hat{Q}_\theta$ can be either the "in-phase" component $\hat{P}$ or the "out-of-phase" component $\hat{Q}$ of the electric field. Therefore by varying the phase of the reference signal, properties of squeezed and anti-squeezed light can be examined.

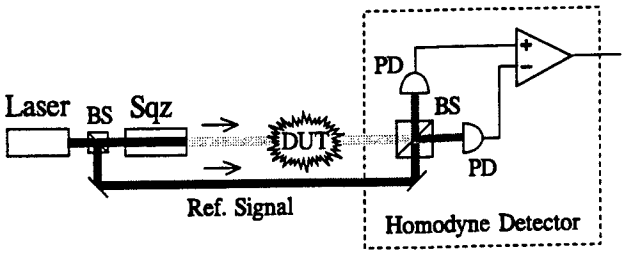


Figure 4.13.1: A system to determine the effect that a device-under-test (DUT) has on the noise in a squeezed beam of electromagnetic energy. The reference signal for the homodyne detector is derived from the laser source.

Next, focus on the homodyne detection system. Assume that the reference signal is represented by a classical amplitude α_{LO} (i.e., the mode has a large amplitude). The symbols $\hat{a}, \hat{a}_1, \hat{a}_2$ refer to the operator amplitudes of the light waves. The following shows that the difference in the number of photons arriving at the two photodetectors is given by

$$\hat{n}_{21} = \sqrt{2}|\alpha_{LO}|\hat{Q}_\theta \qquad (4.13.1)$$

An increase of the signal from the "local oscillator" also increases the size of the detected signal n_{21}; consequently the system provides an amplified version of the quadrature signal Q_θ. The detection system responds only to the frequency of the local oscillator and the desired quadrature component can be selected by using the appropriate value for

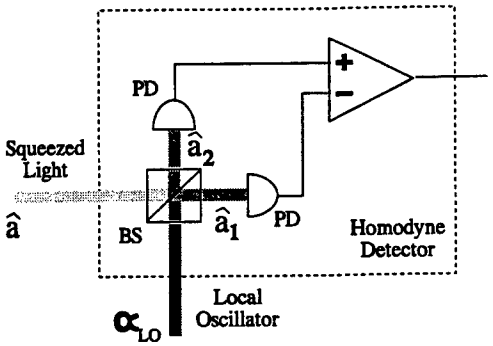


Figure 4.13.2: The homodyne detection system combines squeezed light and the large-amplitude local-oscillator signal at the beam splitter (mixing).

θ. Besides detecting the quadrature component, the variance of the photon-number difference provides the variance of the selected quadrature component which gives a measure of the amount of noise present.

The detectors respond to the power in each beam, which is essentially a measure of the number of photons in each beam. The number of photons in output beam #1 and beam #2, respectively, is given by

$$\hat{n}_1 = \hat{a}_1^+ \hat{a}_1 \qquad \hat{n}_2 = \hat{a}_2^+ \hat{a}_2$$

The amplitudes are defined to be

$$\hat{a}_1 = \frac{1}{\sqrt{2}}(\hat{a} - \alpha_{LO}) \qquad \hat{a}_2 = \frac{1}{\sqrt{2}}(\hat{a} + \alpha_{LO})$$

where the amplitude $\hat{a}_1$ includes a 180° phase shift with the classical amplitude for reflection at the internal interface of the beam splitter. Similar phase shifts are not included for the quantum amplitude $\hat{a}$. The difference in the number of photons incident on the detectors is

$$\begin{aligned}\hat{n}_{21} &= \hat{n}_2 - \hat{n}_1 = \hat{a}_2^+ \hat{a}_2 - \hat{a}_1^+ \hat{a}_1 \\ &= \frac{1}{\sqrt{2}}(\hat{a}^+ + \alpha_{LO}^*)\frac{1}{\sqrt{2}}(\hat{a} + \alpha_{LO}) - \frac{1}{\sqrt{2}}(\hat{a}^+ - \alpha_{LO}^*)\frac{1}{\sqrt{2}}(\hat{a} - \alpha_{LO}) \\ &= \alpha_{LO}\hat{a}^+ + \alpha_{LO}^*\hat{a}\end{aligned}$$

Using the definition for the creation and annihilation operators in terms of the "position and momentum" operators,

$$\hat{a} = \frac{1}{\sqrt{2}}(\hat{Q} + i\hat{P}) \qquad \hat{a}^+ = \frac{1}{\sqrt{2}}(\hat{Q} - i\hat{P})$$

the difference in photon number becomes

$$\hat{n}_{21} = \alpha_{LO}\hat{a}^+ + \alpha_{LO}^*\hat{a} = \frac{\alpha_{LO}^* + \alpha_{LO}}{\sqrt{2}}\hat{Q} + i\frac{\alpha_{LO}^* - \alpha_{LO}}{\sqrt{2}}\hat{P} = \hat{Q}\sqrt{2}\,\mathrm{Re}(\alpha_{LO}) + \hat{P}\sqrt{2}\,\mathrm{Im}(\alpha_{LO})$$

$$= \sqrt{2}|\alpha_{LO}|\left(\hat{Q}\,\mathrm{Cos}\theta + \hat{P}\,\mathrm{Sin}\theta\right) = \sqrt{2}|\alpha_{LO}|\,\hat{Q}_\theta$$

Section 4.14: Production of Quantum EM States

A number of techniques are available for producing Quantum Electromagnetic States. Fock states are number-squeezed states with a number-variance of zero. These higly squeezed number states have not been demonstrated to date. The coherent states are minimum-area uncertainty states. The coherent states can be produced by classical currents and by lasers operating well above threshold. Using low noise current sources to bias light emitting diodes and lasers can produce number squeezed light. Quadrature squeezed light can be generated using four-wave mixing and parametric amplifiers and properly constructed and properly biased semiconductor lasers. A full review of producing quantum EM states must wait for a subsequent book. This section shows that a classical current gives rise to coherent states.

The following shows that a classical current can generate a coherent state (refer to Mandel and Wolf P. 569). Start with the vacuum state $|0\rangle$ for a single mode electric field. The following shows that the time-evolution operator $\hat{U}(t)$ produces a (time-dependent) displacement of the vacuum state. A displaced vacuum state is a coherent state.

The total interaction Hamiltonian is

$$\hat{V}_I(t) = -\int_{vol} d^3r\, \vec{J}(\vec{r},t)\cdot\hat{A}(\vec{r},t) \tag{4.14.1}$$

where $\vec{J}$ is a (real) classical current density, $\hat{A}$ is the single-mode quantized vector potential (in the Coulomb guage) and the subscript "I" indicates the interaction representation. For the interaction representation, recall that the operator caries the trivial time dependence (i.e., the Heisenberg representation when there is no interaction potential) and the wavefunction caries the dynamics of the interaction. Assume the vector potential remains small enough that it does not affect the current density; the current density is therefore assumed to be known. The vector potential has the form

$$\hat{A} = \sqrt{\frac{\hbar}{2\varepsilon_o V\omega}}\,\tilde{e}\left[\hat{b}e^{i\vec{k}\cdot\vec{r}-i\omega t} + \hat{b}^+e^{-i\vec{k}\cdot\vec{r}+i\omega t}\right] \tag{4.14.2}$$

where the creation and annihilation operators in the vector potential are independent of time and $\tilde{e}$ the optical polarization vector. Substituting Equation 4.14.2 into Equation 4.14.1, we obtain

$$\hat{V}_I(t) = -\sqrt{\frac{\pi\hbar}{\varepsilon_o V\omega}}\,\tilde{e}\cdot\left[\hat{b}\,\tilde{J}(k,t)e^{-i\omega t} + \hat{b}^+\tilde{J}^*(k,t)e^{+i\omega t}\right] \tag{4.14.3}$$

where the Fourier transform of the current density is defined by

$$\tilde{J}(\vec{k},t) = \int_{-\infty}^{\infty} d^3r\, \vec{J}(\vec{r},t)\frac{e^{i\vec{k}\cdot\vec{r}}}{\sqrt{2\pi}}$$

Assume that the interaction starts at t=0.

First, as in Chapter 2, the discussion shows that Schrodinger equation for the interaction representation has the form

$$V_I(t)\,\Psi(x,t) = i\hbar\frac{\partial}{\partial t}\Psi(x,t)$$

where the total Hamiltonian accounts for the energy of the free field, free current and the interaction between them.

$$\hat{H} = \hat{H}_o + \hat{V} = \underset{EM}{\hat{H}_{free}} + \underset{current}{\hat{H}_{free}} + \hat{V} = \underset{EM}{\hat{H}_{free}} + \hat{V}$$

The Hamiltonian $\hat{H}_o = \underset{EM}{\hat{H}_{free}} + \underset{curr}{\hat{H}_{free}} = \underset{EM}{\hat{H}_{free}}$ excludes the "free current" since that current is not quantized. The $\hat{V}$ term gives the interaction energy between the current and the fields. The interaction wave vector is

$$|\Psi(t)\rangle = \exp\left(\frac{\hat{H}_o}{i\hbar}t\right)|\Psi_I(t)\rangle$$

Substituting this into Schrodinger's equation

$$\left[\hat{H}_o + \hat{V}(t)\right]|\Psi(t)\rangle = i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle$$

yields

$$\hat{H}_o \exp\left(\frac{\hat{H}_o}{i\hbar}t\right)|\Psi_I(t)\rangle + \hat{V}(t)\exp\left(\frac{\hat{H}_o}{i\hbar}t\right)|\Psi_I(t)\rangle = \hat{H}_o \exp\left(\frac{\hat{H}_o}{i\hbar}t\right)|\Psi_I(t)\rangle + i\hbar \exp\left(\frac{H_o}{i\hbar}t\right)\frac{\partial}{\partial t}|\Psi_I(t)\rangle$$

Using evolution operator for the Heisenberg representation is

$$\hat{u}(t) = \exp\left(\frac{\hat{H}_o t}{i\hbar}\right)$$

and canceling terms from the previous equation, gives

$$\hat{V}_I(t)|\Psi_I(t)\rangle = i\hbar\frac{\partial}{\partial t}|\Psi_I(t)\rangle$$

where

$$\hat{V}_I(t) = \hat{u}^+(t)\hat{V}(t)\hat{u}(t)$$

The time "t" is restricted to be small $t = \Delta t$ so that there is little change in the operators.

The evolution operator in the Interaction representation can be found by writing a formal solution to

$$\hat{V}_I(t)|\Psi_I(t)\rangle = i\hbar\frac{\partial}{\partial t}|\Psi_I(t)\rangle$$

as

$$|\Psi_I(t)\rangle = \hat{U}(t)|\Psi(0)\rangle$$

with

$$\hat{U}(t) = \exp\left[\frac{1}{i\hbar}\int_0^t d\tau\, \hat{V}_I(\tau)\right]$$

Explicitly taking into account $t = \Delta t$, the interaction evolution operator becomes

$$\hat{U}(t) = \exp\left[\frac{\Delta t}{i\hbar}\hat{V}_I(0)\right]$$

The wave function $|\Psi_I(t+\Delta t)\rangle$ can be written based on its value at an infinitesimally earlier time $|\Psi_I(t)\rangle$

$$|\Psi_I(t+\Delta t)\rangle = \hat{U}(t+\Delta t,t)|\Psi_I(t)\rangle$$

where

$$\hat{U}(t+\Delta t,t) = \exp\left[\frac{\Delta t}{i\hbar}\hat{V}_I(t)\right]$$

and $\Delta t = t - t'$. Defining the function

$$f^*(t) = -\frac{i}{\hbar}\sqrt{\frac{\pi\hbar}{\omega\varepsilon_o V}}\,\tilde{e}\cdot\tilde{J}(\vec{k},t)e^{-i\omega t}$$

to substitute into the interaction Hamiltonian to get

$$\hat{V}_I(t) = -\sqrt{\frac{\pi\hbar}{\varepsilon_o V\omega}}\tilde{e}\cdot\left[\hat{b}\,\tilde{J}(k,t)e^{-i\omega t} + \hat{b}^+\tilde{J}^*(k,t)e^{+i\omega t}\right] = -\frac{\hbar}{i}\left[f^*(t)\hat{b} - f(t)\hat{b}^+\right]$$

The evolution operator becomes

$$\hat{U}(t+\Delta t,t) = \exp\left[-\Delta t\, f^*(t)\hat{b} + \Delta t\, f(t)\hat{b}^+\right]$$

This is seen to be a coherent-state displacement operator by comparing it with

$$D(\alpha) = e^{\left(\alpha\hat{b}^+ - \alpha^*\hat{b}\right)}$$

so that

$$\hat{U}(t+\Delta t,t) = D[\Delta t\, f(t)]$$

Therefore the unitary operator $\hat{U}$ causes the vacuum state $|0\rangle$ to evolve into the coherent state

$$|\Delta\alpha\rangle = D[\Delta t\, f(t)]|0\rangle \qquad (4.14.4)$$

(since it is a displaced vacuum).

Equation 4.14.4 can be integrated to find the coherent state as a function of time. Consider repeatedly applying the evolution operator.

$$|\alpha(t_1)\rangle = D[\Delta t\, f(t_1)]|0\rangle$$

$$|\alpha(t_2)\rangle = D[\Delta t\, f(t_2)]|\alpha(t_1)\rangle = D[\Delta t\, f(t_2)]\,D[\Delta t\, f(t_1)]|0\rangle$$

$$\vdots$$

$$|\alpha(t_n)\rangle = \left\{\prod_{i=1}^{n} D[\Delta t\, f(t_i)]\right\}|0\rangle$$

The infinitesimal displacement from the origin $\Delta\alpha(t) = \Delta t\, f(t)$ depends on time. Using the relation

$$\hat{D}(\alpha)\hat{D}(\beta) = e^{\left(\alpha\beta^* - \alpha^*\beta\right)/2}\hat{D}(\alpha+\beta)$$

where $\exp\left\{\left(\alpha\beta^* - \alpha^*\beta\right)/2\right\}$ is a phase factor, we can write

$$|\alpha(t)\rangle = c\,D\left[\sum_{i=1}^{n} f(t_i)\,\Delta t\right]|0\rangle = c\,D\left[\int_0^t d\tau\, f(\tau)\right]|0\rangle = c\left|\int_0^t d\tau\, f(\tau)\right\rangle$$

where "c" is a phase factor.

Chapter 5: Matter-Field Interaction

The equations of motion for matter and electromagnetic (EM) fields are of little value without the possibility of an interaction between them. It is this matter-field interaction that provides for a host of devices and systems such as lasers, LEDs, RF transmitters and the sun. The wavefunctions for the matter and the fields occupy separate Hilbert spaces. The motions of the wavefunctions in their respective spaces become linked through the matter-field interaction. For example, accelerating charge in the matter space produces greater field amplitude in the other (as discussed in Section 4.14).

The density operator provides the most information possible on the state of the combined system (fields and matter). This operator lives in the direct product space consisting of (at minimum) the product of the spaces for the matter and field. The number of spaces can increase depending on the number of degrees of freedom. For example, reservoirs that provide the Langevin (fluctuation) and damping forces each have wavefunction that reside in their own Hilbert space. The combination of the matter, fields and reservoirs is considered to be a complete system.

The motion of the density operator in its direct product space provides information on the transitions that the system makes. An atom, for example, might have an electron in the first excited state but, through the matter-field interaction, it can make a transition to the lower state while exciting one of the optical modes. The *master equation* (a.k.a., the Liouville equation for the density operator) is a first order differential equation that relates the rate of change of the density operator to occupation probabilities, transition rates, and to fluctuation and damping terms. The motion of the density operator can be treated in a fully quantum mechanical way.

The objective of this chapter is to introduce the master equation without solving it for any specific problems. This is in keeping with the direction chosen for this book, namely to summarize the mathematical apparatus and concepts required for future applications. The primary reason for studying the quantum theory to the depth presented in this book is that the density operator method and dynamical equations are general enough to describe all known physical situations. The formalism is easily applied to the interaction of an electromagnetic field as characterized by one of the quantum states (Fock, coherent or squeezed).

The Liouville equation and its solution provide a natural explanation for Homogenous level-broadening due to "loss" mechanisms. If Schrodinger's equation is solved for the complete system then the eigenfunctions (existing in a possibly large direct product space) are exact. Once the complete system collapses into one of the eigenstates, it will remain there. Now, suppose that a smaller system is not complete in the sense that an agent external to the system is capable of interacting with it. Imagine Schrodinger's equation is solved without taking into account the interaction (if the interaction is included then the small system is identical to the complete one). The eigenfunction solutions for the smaller system are not the same as those for the complete system because the Hamiltonian for the small one does not include the interaction energy. If the small system collapses to one of its eigenstates, the perturbing external agent can cause it to make transitions from one eigenstate to another. In particular, external agents interact with the oscillating dipole moment between states. The lifetime of the state is consequently uncertain; the state no longer has an infinite lifetime. As a result of the

Heisenberg uncertainty relation for energy-time, it is clear that the energy of the state is no longer certain. This is equivalent to saying the level is broadened. For this argument, it does not matter whether the external agent is a "loss" (such as losing light through a mirror) or a "gain" mechanism. It only matters that the Hamiltonian of the smaller system is not exact and the agent interferes with the coherence.

The density operator approach for lasers results in the laser rate equations. The rate equations essentially provide all of the information necessary to model the matter-light interaction for laser operation. There are, of course, addition design equations for waveguiding and the electrical-device aspects.

Section 5.1: Introduction to the Liouville Equation for the Density Operator

The master equation (Liouville equation) for the density operator is usually first encountered in a semiclassical setting where the matter is quantized but the electromagnetic field is not. The density operator $\hat{\rho}$ provides information on the occupation and transition probability. The diagonal elements are the probability of finding the system in a given state whereas the off-diagonal terms are related to the induced polarization (refer to Section 2.17). As the density operator changes with time, the probability of occupying a given state must also change; consequently, the rate of change of the diagonal elements of the density operator must be related to the rate of transition. Therefore, to predict the transition rate (related to the gain of a laser) one must know the time rate of change of the density operator. The discussion considers two phenomenological methods of finding the rate equation for the density operator. The first method uses the Hamiltonian for the complete system, including all relevant terms such as the matter-light interaction, electrical and optical pumping, and collisions. The second method uses a restricted Hamiltonian that includes only the matter-light interaction but the probability in the density matrix becomes time-dependent. Subsequent sections demonstrate the purely quantum mechanical route the ends with the same results. The discussion is geared more toward light generation for lasers in this first section.

Topic 5.1.1: The Liouville Equation Using the Full Hamiltonian

Define the total Hamiltonian as

$$\hat{H} = \hat{H} + \hat{H}_{other} \tag{5.1.1}$$

Assume that only stimulated emission and absorption are included in $\hat{H}$ which is defined as

$$\hat{H} = \hat{H}_o + \hat{V}$$

where $\hat{H}_o$ is the Hamiltonian that describes the atom and $\hat{V}$ is the interaction energy that describes the interaction between an applied electromagnetic wave and the atom. The laser gain is related to $\hat{H}$. The term $\hat{H}_{other}$ describes the effects of pumping currents, collisions, spontaneous emission and other terms not described by stimulated emission and absorption. The $\hat{H}_{other}$ Hamiltonian should be divided into separate terms such as

$$\hat{H}_{other} = \hat{H}_{pump} + \hat{H}_{coll} + \hat{H}_{spont} + \ldots$$

These terms are treated phenomenologically in the next several topics; each of them leads to a separate time constant for the system. For simplicity, the single Hamiltonian $\hat{H}_{other}$ is retained which accounts for all the "non-stimulated" terms. There is a single time constant but the discussion indicates how others obtain from the individual Hamiltonians in $\hat{H}_{other}$.

The Liouville equation for the density operator is found by differentiating the density operator with respect to time

$$\hat{\rho}(t) = \sum_{\psi} P_{\psi} |\psi(t)\rangle\langle\psi(t)|$$

As discussed later, this equation assumes that the wave function carries all of the system dynamics and that the probability P_ψ is independent of time. The differentiation provides

$$\begin{aligned}\frac{\partial\hat{\rho}}{\partial t} &= \sum_\psi P_\psi \frac{\partial|\psi\rangle}{\partial t}\langle\psi| + \sum_\psi P_\psi |\psi\rangle\frac{\partial\langle\psi|}{\partial t} \\ &= \sum_\psi P_\psi \frac{\partial|\psi\rangle}{\partial t}\langle\psi| + \sum_\psi P_\psi |\psi\rangle\left\langle\frac{\partial}{\partial t}\psi\right|\end{aligned} \tag{5.1.2}$$

Notice that for the last term, the partial derivative was moved inside the "bra" which is permissible by recalling the definition for the bra of a function

$$\frac{\partial}{\partial t}\langle\psi| = \frac{\partial}{\partial t}\int dV\,\psi^*(\vec{r},t) = \int dV\,\frac{\partial}{\partial t}\psi^*(\vec{r},t)$$

where the integral is treated as an operator. In this case, the projection operator is assumed to act on a time-independent basis vector (i.e., the matrix element is to be calculated). Next, using Schrodinger's equation

$$\hat{H}|\psi\rangle = i\hbar\frac{\partial}{\partial t}|\psi\rangle \quad\rightarrow\quad \frac{\partial|\psi\rangle}{\partial t} = \frac{\hat{H}}{i\hbar}|\psi\rangle$$

The definition of adjoint and the fact that the Hamiltonian is Hermitian provides

$$\left\langle\frac{\partial\psi}{\partial t}\right| = \left\langle\frac{\hat{H}\psi}{i\hbar}\right| = \langle\psi|\left(\frac{\hat{H}}{i\hbar}\right)^+ = \langle\psi|\left(\frac{\hat{H}}{-i\hbar}\right)$$

Inserting these last two results into Equation 5.1.2 provides

$$\begin{aligned}\frac{\partial\hat{\rho}}{\partial t} &= \sum_\psi P_\psi \frac{\partial|\psi\rangle}{\partial t}\langle\psi| + \sum_\psi P_\psi |\psi\rangle\left\langle\frac{\partial}{\partial t}\psi\right| \\ &= \sum_\psi P_\psi \frac{\hat{H}}{i\hbar}|\psi\rangle\langle\psi| - \sum_\psi P_\psi |\psi\rangle\langle\psi|\frac{\hat{H}}{i\hbar} \\ &= \frac{\hat{H}}{i\hbar}\left\{\sum_\psi P_\psi |\psi\rangle\langle\psi|\right\} - \left\{\sum_\psi P_\psi |\psi\rangle\langle\psi|\right\}\frac{\hat{H}}{i\hbar}\end{aligned}$$

Recognizing the terms in braces are the density operator, this last expression gives a commutator

$$\frac{\partial\hat{\rho}}{\partial t} = \frac{1}{i\hbar}\left[\hat{H},\hat{\rho}\right]$$

Finally, inserting the definition for the Hamiltonian from Equation 5.1.1 to obtain Liouville's equation for the density operator

$$\frac{\partial\hat{\rho}}{\partial t} = \frac{1}{i\hbar}\left[\hat{H},\hat{\rho}\right] = \frac{1}{i\hbar}\left[\hat{H}+\hat{H}_{other},\hat{\rho}\right] = \frac{1}{i\hbar}\left[\hat{H},\hat{\rho}\right] + \frac{1}{i\hbar}\left[\hat{H}_{other},\hat{\rho}\right] \tag{5.1.3}$$

Assume that one does not exactly know $\hat{H}_{other}$ and therefore treats it phenomenologically by replacing the corresponding commutator with

$$\left(\frac{\partial\hat{\rho}}{\partial t}\right)_{other} = \frac{1}{i\hbar}\left[\hat{H}_{other},\hat{\rho}\right]$$

Therefore, an alternate form for Liouville's equation for the density operator is

$$\frac{\partial\hat{\rho}}{\partial t} = \frac{1}{i\hbar}\left[\hat{H},\hat{\rho}\right] = \frac{1}{i\hbar}\left[\hat{H}+\hat{H}_{other},\hat{\rho}\right] = \frac{1}{i\hbar}\left[\hat{H},\hat{\rho}\right] + \left(\frac{\partial\hat{\rho}}{\partial t}\right)_{other} \tag{5.1.4}$$

This equation says that the occupation of the individual atomic states changes (the $\frac{\partial\hat{\rho}}{\partial t}$ term) due to the interaction of electromagnetic field with the atom (the $\frac{1}{i\hbar}\left[\hat{H},\hat{\rho}\right]$ term) and due to other sources such as electrical pump currents (the $\left(\frac{\partial\hat{\rho}}{\partial t}\right)_{other} = \frac{1}{i\hbar}\left[\hat{H}_{other},\hat{\rho}\right]$ term). This last term is due to the relaxation of the carrier relaxation that brings the system back to equilibrium once the electromagnetic perturbation is removed. Therefore, it is the last term in Equation 5.1.4 that leads to a relaxation time τ.

Consider the full form for $\hat{H}_{other}$

$$\hat{H}_{other} = \hat{H}_{pump} + \hat{H}_{coll} + \hat{H}_{spont} + \ldots$$

in Equation 5.1.3. The Liouville Equation then has the form

$$\frac{\partial\hat{\rho}}{\partial t} = \frac{1}{i\hbar}\left[\hat{H},\hat{\rho}\right] + \frac{1}{i\hbar}\left[\hat{H}_{pump},\hat{\rho}\right] + \frac{1}{i\hbar}\left[\hat{H}_{coll},\hat{\rho}\right] + \frac{1}{i\hbar}\left[\hat{H}_{spont},\hat{\rho}\right] + \ldots \tag{5.1.5}$$

which can also be written as

$$\frac{\partial\hat{\rho}}{\partial t} = \frac{1}{i\hbar}\left[\hat{H},\hat{\rho}\right] + \left(\frac{\partial\hat{\rho}}{\partial t}\right)_{pump} + \left(\frac{\partial\hat{\rho}}{\partial t}\right)_{coll} + \left(\frac{\partial\hat{\rho}}{\partial t}\right)_{spont} + \ldots \tag{5.1.6}$$

As a result, there can be three or more time-constants because of the "pump", "collision" and "spontaneous emission" Hamiltonians in Equation 5.1.5.

Topic 5.1.2: The Liouville Equation Using a Time-dependent Probability

The second phenomenological method of demonstrating the Liouville equation focuses on the Hamiltonian for the atom and stimulated emission-absorption

$$\hat{H} = \hat{H}_o + \hat{V} \tag{5.1.7}$$

The interaction energy $\hat{V}$ describes the interaction of light and matter to produce stimulated emission and absorption. As before, $\hat{H}_o$ is the atomic Hamiltonian. This time, in the definition for the density operator

$$\hat{\rho}(t) = \sum_{\psi} P_{\psi}(t)\left|\psi(t)\right\rangle\left\langle\psi(t)\right| \tag{5.1.8}$$

the probability is allowed to change with time (private communication with Prof. C.L. Tang at Cornell University). In this approach, $\left|\psi(t)\right\rangle$ is required to satisfy the *restricted* Shrodinger equation

$$\hat{H}\left|\psi(t)\right\rangle = i\hbar\frac{\partial}{\partial t}\left|\psi(t)\right\rangle$$

where $\hat{H} = \hat{H}_o + \hat{V}$. This means that some physical mechanisms capable of changing the occupation probabilities are not included in the Hamiltonian $\hat{H}$ and, consequently, they are not included in $\left|\psi(t)\right\rangle$. Therefore the only alternative is to include these "other"

effects by requiring the probability P_ψ to be time-dependent; the probabilities P_ψ depend on time to account for the possible relaxation effects not included in $|\psi(t)\rangle$. The "other" terms include the effects of the pumping currents, spontaneous emission, and non-radiative decay due to collisions. The wavefunctions describe the atom and the transitions caused by the electromagnetic interaction. Taking the probability as a function of time results in a phenomenological term in the Liouville equation for the density operator.

As an example, consider a bag with 5 atoms. Assume all of the electrons are either in eigenstate $|u_1\rangle$ or $|u_2\rangle$. There are only two possible wavefunctions in the density operator

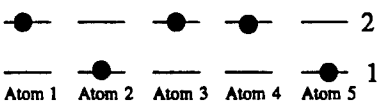


Figure 5.1.1: Five two-level atoms. The eigenstates are labeled "1" and "2".

$$\hat{\rho}(t) = \sum_{S=1}^{2} P_S |u_S\rangle\langle u_S|$$

where "S" takes on values of S=1,2 (assuming the non-coherent case for the density operator whereby linear combinations of the basis vectors are not allowed). Figure 5.1.1 shows that the probability of wavefunction #2 is $P_2 = 3/5$ and the probability of wavefunction #1 is $P_1 = 2/5$. Now suppose all of the optical and electrical pumping is turned-off. If the wavefunctions $|u_S\rangle$ are independent of time (as is the case for basis vectors) then how do the carriers relax to level #1? In particular, what allows the probability of an electron occupying level #2, which is ρ_{22}, to decrease with time? The answer is that the probability P_s must depend on time.

The Liouville equation for the density operator is found by differentiating Equation 5.1.2 with respect to time.

$$\begin{aligned}\frac{\partial\hat{\rho}}{\partial t} &= \sum_\psi \frac{\partial P_\psi}{\partial t}|\psi\rangle\langle\psi| + \sum_\psi P_\psi \frac{\partial|\psi\rangle}{\partial t}\langle\psi| + \sum_\psi P_\psi |\psi\rangle\frac{\partial\langle\psi|}{\partial t} \\ &= \sum_\psi \frac{\partial P_\psi}{\partial t}|\psi\rangle\langle\psi| + \sum_\psi P_\psi \frac{\partial|\psi\rangle}{\partial t}\langle\psi| + \sum_\psi P_\psi |\psi\rangle\left\langle\frac{\partial}{\partial t}\psi\right|\end{aligned} \qquad (5.1.9)$$

Again, using Schrodinger's equation we can write

$$\frac{\partial|\psi\rangle}{\partial t} = \frac{\hat{H}}{i\hbar}|\psi\rangle$$

and also

$$\left\langle\frac{\partial\psi}{\partial t}\right| = \left\langle\frac{\hat{H}\psi}{i\hbar}\right| = \langle\psi|\left(\frac{\hat{H}}{i\hbar}\right)^+ = \langle\psi|\left(\frac{\hat{H}}{-i\hbar}\right)$$

Inserting these last two results into Equation 5.1.9 provides

$$\frac{\partial\hat{\rho}}{\partial t} = \sum_{\psi}\frac{\partial P_{\psi}}{\partial t}|\psi\rangle\langle\psi| + \sum_{\psi}P_{\psi}\frac{\partial|\psi\rangle}{\partial t}\langle\psi| + \sum_{\psi}P_{\psi}|\psi\rangle\left\langle\frac{\partial}{\partial t}\psi\right|$$

$$= \sum_{\psi}\frac{\partial P_{\psi}}{\partial t}|\psi\rangle\langle\psi| + \sum_{\psi}P_{\psi}\frac{\hat{H}}{i\hbar}|\psi\rangle\langle\psi| - \sum_{\psi}P_{\psi}|\psi\rangle\langle\psi|\frac{\hat{H}}{i\hbar}$$

$$= \sum_{\psi}\frac{\partial P_{\psi}}{\partial t}|\psi\rangle\langle\psi| + \frac{\hat{H}}{i\hbar}\left\{\sum_{\psi}P_{\psi}|\psi\rangle\langle\psi|\right\} - \left\{\sum_{\psi}P_{\psi}|\psi\rangle\langle\psi|\right\}\frac{\hat{H}}{i\hbar}$$

Recognizing the terms in braces as the density operator, we can write this last expression as a commutator relation

$$\frac{\partial\hat{\rho}}{\partial t} = \frac{1}{i\hbar}\left[\hat{H},\hat{\rho}\right] + \sum_{\psi}\frac{\partial P_{\psi}}{\partial t}|\psi\rangle\langle\psi| \qquad (5.1.10)$$

which is Liouville's equation for the density operator. The first term is the quantum mechanical interaction. The first term is specialized to the interaction necessary to produce stimulated emission. The last term is assumed to contain the description of collisions, pumping currents, and spontaneous emission. Therefore the commutator will be related to the quantum mechanical gain. The last term is the carrier relaxation term that brings a system back to steady-state once the electromagnetic perturbation is removed. Notice that the wave function does change in Equation 5.1.10 due to a change in the probability (this is because the density operator itself is defined in terms of the wave function).

Topic 5.1.3: The Relaxation Term

This section considers the statistical effects of relaxation on the density operator by working with "other" terms in the Liouville equation. A comparison of Equations 5.1.10, 5.1.3 and 5.1.4 shows

$$\left(\frac{\partial\hat{\rho}}{\partial t}\right)_{other} = \frac{1}{i\hbar}\left[\hat{H}_{other},\hat{\rho}\right] = \sum_{\psi}\frac{\partial P_{\psi}}{\partial t}|\psi\rangle\langle\psi| \qquad (5.1.10)$$

As already mentioned, the first term on the right hand side of

$$\frac{\partial\hat{\rho}}{\partial t} = \frac{1}{i\hbar}\left[\hat{H},\hat{\rho}\right] + \left(\frac{\partial\hat{\rho}}{\partial t}\right)_{other}$$

represents (in part) the quantum mechanical light-matter interaction to produce stimulated emission or absorption while the second term represents the "other" influences that can change the occupation number. Figure 5.1.2 shows an initial wave function $|\psi(0)\rangle$ that has been disturbed from its steady-state position in Hilbert space. Assuming that the electromagnetic perturbation is turned-off, the wave function must decay to some steady-state position which means that the occupation probability for atomic levels must decay to steady state

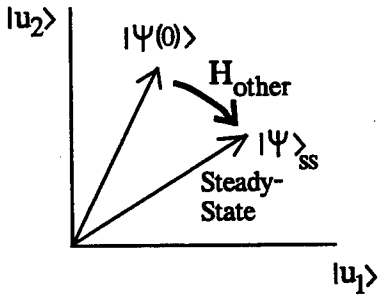


Figure 5.1.2: $\hat{H}_{other}$ causes a disturbed wave function to relax to steady state.

values. Figure 5.1.2 shows that the Hamiltonian $\hat{H}_{other}$ is responsible for the decay.

Equation 5.1.10 requires that a change in the initial wave function depicted Figure 5.1.2 because of the term involving $\frac{\partial P_\psi}{\partial t}$. Assume that the rate of decay of the probability is proportional to the difference between the probability at time t, denoted by $P_\psi(t)$, and the steady state value of the probability which is denoted by $\overline{P_\psi}$. Let τ be the rate of decay that primarily describes the effects of spontaneous emission, nonradiative decay and electrical pumping. The change in the probability can be taken as

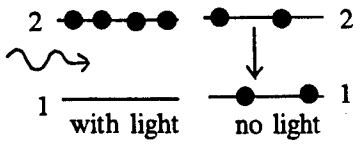


Figure 5.1.3: Cartoon showing that when optical pumping is removed, the atoms relax.

$$\frac{\partial P_\psi}{\partial t} = -\frac{P_\psi(t) - \overline{P_\psi}}{\tau} \tag{5.1.11}$$

This essentially gives the rate of decay of the carriers from one band to the other, for example without electrical pumping, when any impressed electromagnetic field is removed. For example, Figure 5.1.3 shows two-level atoms; the atoms decay into the lower-level to achieve steady state which means that ρ_{22} must also be decaying (since ρ_{22} is the probability that the second level is occupied). Equation 5.1.11 is a simple first-order differential equation and is easy to solve.

$$P_\psi(t) = \overline{P}_\psi + \left[P_\psi(0) - \overline{P}_\psi\right]e^{-t/\tau}$$

Substituting Equation 5.1.11 into the Liouville equation

$$\frac{\partial \hat{\rho}}{\partial t} = \frac{1}{i\hbar}\left[\hat{H}, \hat{\rho}\right] + \sum_\psi \frac{\partial P_\psi}{\partial t}|\psi\rangle\langle\psi|$$

we find

$$\frac{\partial \hat{\rho}}{\partial t} = \frac{1}{i\hbar}\left[\hat{H}, \hat{\rho}\right] - \sum_\psi \frac{P_\psi(t) - \overline{P_\psi}}{\tau}|\psi\rangle\langle\psi|$$

which gives an alternate form of the Liouville equation

$$\frac{\partial \hat{\rho}}{\partial t} = \frac{1}{i\hbar}\left[\hat{H}, \hat{\rho}\right] - \frac{\hat{\rho}(t) - \bar{\rho}}{\tau} \tag{5.1.12}$$

The last term in Equation 5.1.12 is a phenomenological term that provides for carrier relaxation effects. Yariv's Quantum Electronics book solves the equation for lasers.

Section 5.2: Jaynes-Cummings' Model for Matter-Field Interactions

The Jaynes-Cummings model treats the interaction between a two-level radiator/absorber (atom) and an electromagnetic (EM) field. In particular, it considers the interaction between an atom and an optical field for a single optical mode. The Hamiltonian for the atom, light and interaction is given a second quantized form. The more general model of a two-level atom interacting with a large number of modes is considered. Subsequent sections incorporate the effects of quantum mechanical reservoirs into the model. When reservoirs are included in the theory, the complete system is essentially divided into a smaller system and the reservoir. As discussed in the introduction to Chapter 5, the smaller system responds to the influence of the reservoirs which, for lasers, account for optical loss, pumping, collisions (etc). The reservoirs give rise to fluctuations and relation (the "other" part of the Liouville equation). In this section, we consider the smaller system.

The system itself consists of two parts, namely the atom and the field. In the absence of interaction between the atom and field, the atom exists in its own Hilbert space while the light exists in its own. Only when the atom and light interact is there a connection between the two Hilbert spaces.

Topic 5.2.1: The Atomic Hamiltonian

The Hamiltonian for the system consists of an atom, field and interaction Hamiltonian. Consider the atomic Hamiltonian denoted by $\hat{H}_a$. Assume a two-level atom with energy eigenstates denoted by $\{|n\rangle = |E_n\rangle \text{ for } n = 1,2\}$. The atomic Hamiltonian is written in terms of projection operators as if Schrodinger's equation is already solved. All of the essential information is embodied in this abstract representation; there is no need to have the explicit functional form of the wave functions and operators. The closure relation for the two-dimensional space

$$\sum_{n=1}^{2} |n\rangle\langle n| = 1$$

is used to rewrite the atomic Hamiltonian as

$$\hat{H}_a = \hat{H}_a \sum_{n=1}^{2} |n\rangle\langle n| = \sum_{n=1}^{2} \hat{H}_a |n\rangle\langle n| = \sum_{n=1}^{2} E_n |n\rangle\langle n| = E_1 |1\rangle\langle 1| + E_2 |2\rangle\langle 2| \qquad (5.2.1)$$

The Hamiltonian consists of a linear sum of linear operators (termed Pauli oprators because of their similarity with spin operators). The Pauli operators can be given special symbols

$$\hat{\sigma}_{ij} = |i\rangle\langle j|$$

and

$$\hat{\sigma}_z = \hat{\sigma}_{22} - \hat{\sigma}_{11} = |2\rangle\langle 2| - |1\rangle\langle 1|$$
$$\hat{\sigma}_y = i(\hat{\sigma}_{12} - \hat{\sigma}_{21})$$
$$\hat{\sigma}_x = \hat{\sigma}_{12} + \hat{\sigma}_{21}$$

The operator $\hat{\sigma}_z$ essentially gives the difference in population (i.e. probability) between the second and first energy levels.

Example: Suppose

$$|\psi\rangle = \frac{1}{\sqrt{2}}|1\rangle + \frac{1}{\sqrt{2}}|2\rangle$$

which says that a given atom is 50% in the first energy level and 50% in the second. The average difference in population is

$$\langle\psi|\hat{\sigma}_z|\psi\rangle = 0$$

Sometimes it is simpler to compute quantities (such as commutators) using the matrix representation of the Pauli operators. The matrix equivalent of the Pauli operators are easily seen to be

$$\underline{\sigma}_z = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \quad \underline{\sigma}_y = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \quad \underline{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

For example, the equation $\langle\psi|\hat{\sigma}_z|\psi\rangle = 0$ (in the previous example) takes the form

$$\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}^+ \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} = 0$$

The raising and lowering operators are perhaps the most important operators for the atomic Hamiltonian. The raising and lowering operators are defined respectively by

$$\hat{\sigma}^+ = |2\rangle\langle 1| \quad \hat{\sigma}^- = |1\rangle\langle 2| \tag{5.2.2}$$

For example, the raising operator promotes an electron from the lower to the higher level

$$\hat{\sigma}^+|1\rangle = |2\rangle \quad \hat{\sigma}^+|2\rangle = 0$$

The raising and lowering operators have a matrix representation

$$\underline{\sigma}^+ = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad \underline{\sigma}^- = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

The vectors

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

represent an electron in states $|2\rangle$ and $|1\rangle$ respectively. The commutation relations are easy to calculate using either the operator or matrix form.

$$\left[\hat{\sigma}^-, \hat{\sigma}^+\right] = -\hat{\sigma}_z$$

As an important note, these operator are termed "raising and lowering" even though they are not necessarily related to a harmonic oscillator. The basis vector expansion in 5.2.2 differs from that of the harmonic oscillator ladder operators. The multiplicative constant in the equation just following (5.2.2) is also different.

The atomic Hamiltonian is most conveniently written as an operator that measures the difference in population.

$$\hat{H}_a = E_1|1\rangle\langle 1| + E_2|2\rangle\langle 2| = \frac{1}{2}(E_2 - E_1)\left(|2\rangle\langle 2| - |1\rangle\langle 1|\right) + \frac{1}{2}(E_2 + E_1) = \frac{E_2 - E_1}{2}\hat{\sigma}_z + \frac{E_2 + E_1}{2}$$

The difference in energy between levels is usually defined in terms of an atomic resonance frequency ω_o as

$$\hbar\omega_o = E_2 - E_1$$

The atomic Hamiltonian is then

$$\hat{H}_a = \frac{1}{2}\hbar\omega_o\hat{\sigma}_z + \frac{E_2 + E_1}{2}$$

The energy scale for the atom can be reset so that the zero-energy lies midway between E_2 and E_1 which makes the last term in $\hat{H}_a$ equal to zero. The last term (whether it's zero or not) has no effect on the rate of change of a density operator (for example) since that last term commutes with all other operators (and c-numbers). If $\hat{\rho}_a$ is the atomic density operator then

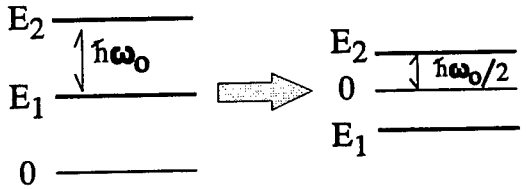


Figure 5.2.1: Redefining the zero of energy for the atom.

$$\frac{\partial\hat{\rho}_a}{\partial t} = \frac{1}{i\hbar}\left[\hat{H}_a, \hat{\rho}_a\right] = \frac{1}{i\hbar}\left[\frac{1}{2}\hbar\omega_o\hat{\sigma}_z + \frac{E_2 + E_1}{2}, \hat{\rho}_a\right] = \frac{1}{i\hbar}\left[\frac{1}{2}\hbar\omega_o\hat{\sigma}_z, \hat{\rho}_a\right]$$

which is obviously independent of the additive c-number term in the atomic Hamiltonian. Either by resetting the energy scale or for the reason of computing rates of change, the added c-number constant term is typically dropped from the Hamiltonian. The final version of the atomic Hamiltonian is

$$\hat{H}_a = \frac{1}{2}\hbar\omega_o\hat{\sigma}_z \tag{5.2.3}$$

Topic 5.2.2: The Free-Field Hamiltonian

The Hamiltonian for the free electromagnetic field is discussed in Chapter 3. The classical Hamiltonian

$$H_f = \int_V dV\left(\frac{\varepsilon_o}{2}\vec{E}\cdot\vec{E} + \frac{1}{2\mu_o}\vec{B}\cdot\vec{B}\right)$$

is replaced by the quantum mechanical one

$$\hat{H}_f = \sum_{\vec{k}s}\hbar\omega_k\left(\hat{b}^+_{\vec{k}s}\hat{b}_{\vec{k}s} + \frac{1}{2}\right) = \sum_{\vec{k}s}\hbar\omega_k\left(\hat{N}_{\vec{k}s} + \frac{1}{2}\right)$$

where $\hat{b}^+_{\vec{k},s}$, $\hat{b}_{\vec{k},s}$ are the creation and annihilation operators, respectively. Recall that the "creation operator" creates a particle in the electromagnetic mode represented by $\vec{k}$ with one of two polarizations represented by "s". The creation and annihilation operators satisfy the commutation relations

$$\left[\hat{b}_{\vec{k},s}, \hat{b}^+_{\vec{K},S}\right] = \delta_{\vec{k},\vec{K}}\delta_{s,S} \tag{5.2.4}$$

The electromagnetic creation and annihilation operators (in either the Schrodinger or the interaction representation) commute for all times with the atomic raising and lowering operators (in either the Schrodinger or the interaction representation). The Heisenberg representation is another story. The Heisenberg representation mixes *all* of the dynamics

into the operators including the interaction between the atom and field. As a result, the Heisenberg operators do not necessarily commute with themselves at different times.

Topic 5.2.3: The Interaction Hamiltonian

The interaction Hamiltonian links the atomic subsystem with the EM subsystem. It is the interaction Hamiltonian that correlates the motion of the EM wave functions and the atomic wave functions in their respective Hilbert spaces. It is the interaction Hamiltonian that drives the motion of the wave functions in the interaction representation.

The basic requirement is that a unit of energy removed from the field must appear in the atom and vice versa. Therefore we expect to find terms of the form

$$\hat{b}^{-}\hat{\sigma}^{+} \quad \text{and} \quad \hat{b}^{+}\hat{\sigma}^{-}$$

These terms link the two Hilbert spaces (i.e., link the two subspaces in the larger direct product space). For a *single* optical mode and a two-level atom, the basis vectors in the direct product space have the form

$$\left\{ |a,n\rangle = |a\rangle|n\rangle \quad a = 1,2 \quad n = 0,1,2\ldots \right\}$$

where "a" stands for the atomic state and "n" stands for the number of photons in the electromagnetic mode (assuming Fock states for the EM basis set). For example,

$$\hat{b}^{-}\hat{\sigma}^{+}|a=1,n=5\rangle = \hat{\sigma}^{+}|a=1\rangle\hat{b}^{-}|n=5\rangle = |a=2\rangle\sqrt{5}|4\rangle$$

A general wave vector in the direct product space has the form

$$|\psi(t)\rangle = \sum_{a,n}\beta_{an}(t)\,|a\rangle|n\rangle$$

For *multiple* optical modes, the interaction Hamiltonian $\hat{H}_{af}$ to contains terms such as

$$\sum_{\vec{k},s}\hat{\sigma}^{-}\hat{b}^{+}_{\vec{k}s} \qquad \text{and} \qquad \sum_{\vec{k},s}\hat{\sigma}^{+}\hat{b}^{-}_{\vec{k}s}$$

The basis vectors in the direct product space consist of the direct product of the atomic basis set with the EM Fock set (for example)

$$\left\{ |a\rangle|m_1,m_2\ldots\rangle = |a\rangle|\{m\}\rangle \quad a = 1,2 \quad m_i = 0,1,\ldots \right\}$$

where m_i is the number of photons in the ith mode.

We can find the second quantized form of the Hamiltonian by starting with

$$\hat{H}_{af} = \hat{\mu}\cdot\hat{E}$$

where $\hat{\mu}$ is the dipole moment operator and $\hat{E}$ is the electric field operator. Both operators are vectors in the physical three-dimensional space. The dipole moment operator can be written as a basis vector expansion

$$\hat{\mu} = \left(\sum_{i=1}^{2}|i\rangle\langle i|\right)\hat{\mu}\left(\sum_{j=1}^{2}|j\rangle\langle j|\right) = \sum_{ij}\vec{\mu}_{ij}\,|i\rangle\langle j|$$

Typically, The atom is assumed not to have a permanent dipole moment $\mu_{ii} = 0$ and that the induced dipole moment has the property that $\mu_{12} = \mu_{21}$. The dipole operator reduces to

$$\hat{\mu} = \vec{\mu}_{12}\left[\,|1\rangle\langle 2| + |2\rangle\langle 1|\,\right] = \left(\hat{\sigma}^{-} + \hat{\sigma}^{+}\right)\vec{\mu}_{12}$$

where the physical size of the dipole is assumed small compared to the electromagnetic wavelength (dipole approximation) and where

$$\vec{\mu}_{12} = \int d^3\vec{r}\; u_1^*(\vec{r})(-e\vec{r})\,u_2(\vec{r})$$

The electric field operator is

$$\hat{E}(\vec{r},t) = \sum_{\vec{k}s}\sqrt{\frac{\hbar\omega_k}{2\varepsilon_o}}\left[\hat{b}_{\vec{k}s}(t)\,f_{\vec{k}}(\vec{r}) + \hat{b}^+_{\vec{k}s}(t)\,f^*_{\vec{k}}(\vec{r})\right]\tilde{e}_{\vec{k}s}$$

where $\tilde{e}_{\vec{k}s}$ is the polarization vector. The mode functions $f_{\vec{k}}$ satisfy the classical Maxwell equations with specified boundary conditions (see Topics 3.3.2 and 3.5.3) and they are normalized according to

$$\int_V dV\, f^*_{\vec{k}}(\vec{r})\, f_{\vec{K}}(\vec{r}) = \delta_{\vec{k}\vec{K}}$$

For plane waves with periodic boundary conditions, the mode function is

$$f_{\vec{k}}(\vec{r}) = \frac{e^{i\vec{k}\cdot\vec{r}}}{\sqrt{V}}$$

Example: Orthonormality of "f" for plane waves. The orthonormality can easily be shown by considering the integral for the 1-D case

$$I = \int_0^L dx\, f^*_k(x)\, f_K(x) = \frac{1}{L}\int_0^L dx\, e^{i(k-K)x}$$

where, for periodic boundary conditions, $k = 2\pi n/L$. If k=K then the integral becomes

$$I = \frac{1}{L}\int_0^L dx\; 1 = 1$$

For $k \neq K$, the integral is

$$I = \frac{1}{L}\int_0^L dx\; \exp\left\{i\frac{2\pi(n-m)}{L}x\right\} = \frac{e^{2\pi i(n-m)} - 1}{2\pi i(n-m)} = 0$$

where the last step follows because (n-m) is an integer. Putting the two results together gives the orthonormality relation.

The interaction Hamiltonian can be rewritten by combining the operator expressions for the dipole moment and the electric field.

$$\hat{H}_{af} = \hat{\mu}\cdot\hat{E}(\vec{r},t) = \sum_{\vec{k}s}\sqrt{\frac{\hbar\omega_k}{2\varepsilon_o}}\left(\hat{\mu}_{12}\cdot\tilde{e}_{\vec{k}s}\right)\left[\hat{\sigma}^-(t) + \hat{\sigma}^+(t)\right]\left[\hat{b}_{\vec{k}s}(t)\,f_{\vec{k}}(\vec{r}) + \hat{b}^+_{\vec{k}s}(t)\,f^*_{\vec{k}}(\vec{r})\right] \quad (5.2.5)$$

where the raising, lowering, creation and annihilation operators are all written in the interaction representation (refer to the following topics). The rotating wave approximation (RWA) allows us to drop the terms

$$\hat{\sigma}^-\hat{b} \text{ and } \hat{\sigma}^+\hat{b}^+$$

The rotating wave approximation is equivalent to dropping terms that do not conserve energy. For example, $\hat{\sigma}^-\hat{b}$ removes a photon from the EM field and also a unit of energy from the atom without placing the extra energy anywhere. The RWA is usually

associated with an integral over time; this turns out to be the case when we calculate the integral of the rate of change of the density operator. The final form of the interaction Hamiltonian is

$$\hat{H}_{af} = \hat{\mu}\cdot\hat{E}(\vec{r},t) = \sum_{\vec{k}s}\sqrt{\frac{\hbar\omega_k}{2\varepsilon_o}}\left(\hat{\mu}_{12}\cdot\vec{e}_{\vec{k}s}\right)\left[\hat{\sigma}^+(t)\,\hat{b}_{\vec{k}s}(t)\,f_{\vec{k}}(\vec{r}) + \hat{\sigma}^-(t)\,\hat{b}^+_{\vec{k}s}(t)\,f^*_{\vec{k}}(\vec{r})\right]$$

Topic 5.2.4: Atomic and Interaction Hamiltonians using Fermion Operators

Some books (e.g., Haken) write the atomic and interaction Hamiltonian in terms of Fermion creation and annihilation operators. Rather than the two vectors used for the atomic states, there are now the following three Fermion Fock states.

$|0,0\rangle$ is the vacuum state

$|1,0\rangle$ specifies an electron in the lowest energy level

$|0,1\rangle$ specifies an electron in the highest energy level

The Fermion creation $\hat{f}_n^+$ and annihilation $\hat{f}_n^-$ operators add or subtract a particle from energy level E_n, respectively. These operators have the following effect.

$$\hat{f}_1^+|0,0\rangle = |1,0\rangle \qquad \hat{f}_1^+|1,0\rangle = 0$$

with similar results for $\hat{f}_2^+$ etc. These operators satisfy the anticommutation relations which allow only one electron per state (Pauli exclusion principle).

$$\left\{\hat{f}_a^-,\hat{f}_b^+\right\} = \delta_{ab} \qquad \left\{\hat{f}_a^-,\hat{f}_b^-\right\} = 0 \qquad \left\{\hat{f}_a^+,\hat{f}_b^+\right\} = 0$$

where the anticommutator is defined by

$$\left\{\hat{A},\hat{B}\right\} = \hat{A}\hat{B} + \hat{B}\hat{A}$$

For example, the anticommutation relation

$$\left\{\hat{f}_a^+,\hat{f}_b^+\right\} = 0$$

yields the Pauli exclusion principle

$$\hat{f}_a^+\hat{f}_a^+|0\rangle = \frac{1}{2}2\hat{f}_a^+\hat{f}_a^+|0\rangle = \frac{1}{2}\left\{\hat{f}_a^+,\hat{f}_a^+\right\}|0\rangle = \frac{1}{2}0|0\rangle = 0$$

So it is not possible to create two particles in a single energy state. The atomic raising and lowering operators are replaced as follows

$$\hat{\sigma}^+ \to \hat{f}_2^+\hat{f}_1^- \quad \text{and} \quad \hat{\sigma}^- \to \hat{f}_1^+\hat{f}_2^-$$

This book uses the raising and lowering operators rather than the creation and annihilation operators.

Section 5.3: The Interaction Representation for the Jaynes-Cummings' Model

The master equation for the density operator is usually presented in the interaction representation. As discussed in Chapter 2, the interaction representation is characterized by wavefunctions and density operators that move through Hilbert space as a result of an interaction Hamiltonian (that describes the interaction energy). The ladder, creation and annihilation operators assume a simple exponential time dependence. This Section discusses the interaction representation for the wave functions, operators and Hamiltonians.

Topic 5.3.1: Atomic Creation and Annihilation Operators

The remainder of Chapter 5 is concerned with the interaction representation for both the atom and the electromagnetic field. Consider the atom first. Chapter 2 shows that the wavefunction $|\psi\rangle$ in the Schrodinger representation is related to the wave function $|\breve{\psi}\rangle$ in the *interaction* representation by

$$|\psi\rangle = \exp\left\{\frac{\hat{H}_a t}{i\hbar}\right\}|\breve{\psi}\rangle \tag{5.3.1}$$

The operator

$$\hat{u}_a = \exp\left\{\frac{\hat{H}_a t}{i\hbar}\right\}$$

is unitary.

The interaction representation of the lowering operator $\breve{\sigma}^-$ is found by requiring expectation values for the Schrodinger and Interaction representation to agree.

$$\langle\psi|\hat{\sigma}^-|\psi\rangle = \langle\breve{\psi}|\hat{u}_a^+\hat{\sigma}^-\hat{u}_a|\breve{\psi}\rangle = \langle\breve{\psi}|\breve{\sigma}^-|\breve{\psi}\rangle$$

so that

$$\breve{\sigma}^- = \hat{u}_a^+\hat{\sigma}^-\hat{u}_a \tag{5.3.2}$$

The operator expansion theorem provides the explicit time-dependent form of the interaction lowering operator. The operator expansion theorem is

$$e^{\hat{A}}\hat{B}e^{-\hat{A}} = \hat{B} + [\hat{A},\hat{B}] + \frac{1}{2!}[\hat{A},[\hat{A},\hat{B}]] + \ldots$$

The interaction representation of the lowering operator is

$$\breve{\sigma}^- = \hat{u}_a^+\hat{\sigma}^-\hat{u}_a = \exp\left\{-\frac{\hat{H}_a t}{i\hbar}\right\}\hat{\sigma}^-\exp\left\{\frac{\hat{H}_a t}{i\hbar}\right\} = \hat{\sigma}^- - \left[\frac{\hat{H}_a t}{i\hbar},\hat{\sigma}^-\right] + \ldots$$

Using Equation 5.2.3, the commutator becomes

$$[\hat{H}_a,\hat{\sigma}^-] = \frac{\hbar\omega_o}{2}[\hat{\sigma}_z,\hat{\sigma}^-] \equiv \frac{\hbar\omega_o}{2}\left\{\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix} - \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}\right\} = \hbar\omega_o\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix} = -\hbar\omega_o\hat{\sigma}^-$$

Therefore, the interaction representation for the lowering operator is

$$\breve{\sigma}^-(t) = \hat{\sigma}^- e^{-i\omega_o t} \tag{5.3.3}$$

where $\hat{\sigma}^-$ is the Schrodinger representation for the atomic lower operator. The adjoint of Equation 5.3.3 provides

$$\breve{\sigma}^{+}(t) = \hat{\sigma}^{+}\, e^{+i\omega_o t}$$

The Hamiltonian has the same form in either the interaction or Schrodinger representation

$$\breve{H}_a = \hat{u}_a^{+}\hat{H}_a\hat{u}_a = e^{-\frac{\hat{H}_a t}{i\hbar}}\hat{H}_a e^{\frac{\hat{H}_a t}{i\hbar}} = \hat{H}_a e^{-\frac{\hat{H}_a t}{i\hbar}} e^{\frac{\hat{H}_a t}{i\hbar}} = \hat{H}_a$$

where the third equality follows since an operator always commutes with a function of that operator.

Topic 5.3.2: The Boson Creation and Annihilation Operators

We can find the interaction representation variables for the electromagnetic operators. The EM-field evolution operator is

$$\hat{u}_f = \exp\left\{\frac{\hat{H}_f t}{i\hbar}\right\}$$

where

$$\hat{H}_f = \sum_{\bar{k}s}\hbar\omega_{\bar{k}}\left(\hat{b}_{\bar{k}s}^{+}\hat{b}_{\bar{k}s} + \frac{1}{2}\right)$$

and the creation/annihilation operators in this Hamiltonian are independent of time. The operator expansion theorem

$$e^{\hat{A}}\hat{B}e^{-\hat{A}} = \hat{B} + \left[\hat{A},\hat{B}\right] + \frac{1}{2!}\left[\hat{A},\left[\hat{A},\hat{B}\right]\right] + \ldots$$

provides the interaction representation of the EM-field creation and annihilation operators

$$\breve{b}_{\bar{k}s}(t) = \hat{u}_f^{+}\hat{b}_{\bar{k}s}\hat{u}_f = \exp\left\{-\frac{\hat{H}_f t}{i\hbar}\right\}\hat{b}_{\bar{k}s}\exp\left\{\frac{\hat{H}_f t}{i\hbar}\right\} = \hat{b}_{\bar{k}s} + \frac{it}{\hbar}\left[\sum_{\bar{k}'s'}\hbar\omega_{\bar{k}}\left(\hat{b}_{\bar{k}'s'}^{+}\hat{b}_{\bar{k}'s'} + \frac{1}{2}\right), \hat{b}_{\bar{k}s}\right] + \ldots$$

The commutator has reduces to

$$\left[\sum_{\bar{k}'s'}\hbar\omega_{\bar{k}'}\left(\hat{b}_{\bar{k}'s'}^{+}\hat{b}_{\bar{k}'s'} + \frac{1}{2}\right), \hat{b}_{\bar{k}s}\right] = \sum_{\bar{k}'s'}\hbar\omega_{\bar{k}'}\left[\hat{b}_{\bar{k}'s'}^{+}\hat{b}_{\bar{k}'s'} + \frac{1}{2}, \hat{b}_{\bar{k}s}\right] = \sum_{\bar{k}'s'}\hbar\omega_{\bar{k}'}\left[\hat{b}_{\bar{k}'s'}^{+}, \hat{b}_{\bar{k}s}\right]\hat{b}_{\bar{k}'s'} = -\hbar\omega_{\bar{k}}\hat{b}_{\bar{k}s}$$

where Equation 5.2.3 is used for the commutator. Working out all of the commutators in the expression for $\breve{b}_{\bar{k}s}$ and collecting terms provides

$$\breve{b}_{\bar{k}s}(t) = \hat{b}_{\bar{k}s}\, e^{-i\omega_{\bar{k}}t}$$

The adjoint of this expression provides the interaction representation of the boson creation operator

$$\breve{b}_{\bar{k}s}^{+}(t) = \hat{b}_{\bar{k}s}^{+}\, e^{-i\omega_{\bar{k}}t}$$

where

$$\hat{b}_{\bar{k}s} = \hat{b}_{\bar{k}s}(0) \quad \text{and} \quad \hat{b}_{\bar{k}s}^{+} = \hat{b}_{\bar{k}s}^{+}(0)$$

The Hamiltonian for the free fields has the same form in either the Schrodinger or interaction representation

$$\breve{H}_f = \sum_{\bar{k}s}\hbar\omega_{\bar{k}}\left(\hat{b}_{\bar{k}s}^{+}(t)\,\hat{b}_{\bar{k}s}(t) + \frac{1}{2}\right) = \sum_{\bar{k}s}\hbar\omega_{\bar{k}}\left(\hat{b}_{\bar{k}s}^{+}e^{-i\omega_{\bar{k}}t}\,\hat{b}_{\bar{k}s}e^{+i\omega_{\bar{k}}t} + \frac{1}{2}\right) = \sum_{\bar{k}s}\hbar\omega_{\bar{k}}\left(\hat{b}_{\bar{k}s}^{+}\,\hat{b}_{\bar{k}s} + \frac{1}{2}\right) = \hat{H}_f$$

Topic 5.3.3: Interaction Representation of the Density Operators

We start the discussion of the interaction representation of the density operator with the atom density operator. Let $|\psi\rangle$ be a Schrodinger wave function for an isolated atom. Let $|\breve{\psi}\rangle$ be the interaction representation of the same wave function. The two wavefunctions are related by

$$|\psi\rangle = \exp\left\{\frac{\hat{H}_a t}{i\hbar}\right\}|\breve{\psi}\rangle \tag{5.3.4}$$

with

$$\hat{u}_a = \exp\left\{\frac{\hat{H}_a t}{i\hbar}\right\} \qquad \text{and} \qquad \hat{H}_a = \frac{1}{2}\hbar\omega_o\hat{\sigma}_z$$

a unitary operator. The Schrodinger atomic density operator ρ_a

$$\rho_a = \sum_\psi |\psi\rangle P_\psi \langle\psi| \tag{5.3.5}$$

obtains from the interaction density operator $\breve{\rho}_a$ by substituting Equation 5.3.4 into Equation 5.3.5

$$\rho_a = \sum_\psi |\psi\rangle P_\psi \langle\psi| = \sum_\psi \left\{\exp\left(\frac{\hat{H}_a t}{i\hbar}\right)|\breve{\psi}\rangle\right\} P_\psi \left\{\exp\left(\frac{\hat{H}_a t}{i\hbar}\right)|\breve{\psi}\rangle\right\}^+$$

This equation simplifies upon use of the Hermitian property of the atomic Hamiltonian $\hat{H}_a = \hat{H}_a^+$ to get

$$\hat{\rho}_a = \exp\left(\frac{\hat{H}_a t}{i\hbar}\right)\left\{\sum_\psi |\breve{\psi}\rangle P_\psi \langle\breve{\psi}|\right\}\exp\left(-\frac{\hat{H}_a t}{i\hbar}\right) = \hat{u}_a\,\breve{\rho}_a\hat{u}_a^+$$

The final form of the interaction density operator is

$$\breve{\rho}_a = \hat{u}_a^+\hat{\rho}_a\hat{u}_a \tag{5.3.6}$$

The free field density operator is found similarly to be

$$\breve{\rho}_f = \hat{u}_f^+\hat{\rho}_f\hat{u}_f$$

with

$$\hat{u}_f = \exp\left\{\frac{\hat{H}_f t}{i\hbar}\right\} \qquad \text{and} \qquad \hat{H}_f = \sum_{\bar{k}s}\hbar\omega_{\bar{k}}\left(\hat{b}_{\bar{k}s}^+\hat{b}_{\bar{k}s} + \frac{1}{2}\right)$$

We could have alternatively obtained Equation 5.3.6 by following a procedure similar to that for Equation 5.3.2.

The next several sections are primarily concerned with the density operator in a direct product space and how it is related to the density operators for the atom and EM fields. The maser equation (i.e., rate equation) for the density operator is obtained in the interaction representation. For this reason, it is important to make a few comments on the interaction representation of the density operator for the direct product space.

The wave functions used in the Jaynes-Cummings' model reside in a direct product space $V_a \otimes V_f$ where V_a and V_f are the Hilbert spaces for the atom and the field respectively. For a *single* optical mode, the general wave function in the direct product space has the form

$$|\Psi\rangle = \sum_{nm} \beta_{nm}(t)\, |n\rangle|m\rangle$$

where $|n\rangle$ is the atomic state (n=1,2) and $|m\rangle$ is the Fock electromagnetic state (m=0,1,...). The density operator is given by

$$\hat{\rho} = \sum_{\Psi} |\Psi\rangle P_{\Psi} \langle\Psi|$$

Only under special circumstances (e.g., without interaction between the atomic and optical subsystems) can the density operator be written as the direct product $\hat{\rho} = \hat{\rho}_a \hat{\rho}_f$. Usually it is assumed that just prior to initiating the interaction between the EM fields and matter (t=0), that the density operator can be factored according to

$$\hat{\rho}(0) = \hat{\rho}_a(0)\, \hat{\rho}_f(0)$$

However, it is always possible to write

$$\hat{\rho}_a = \mathrm{Tr}_f(\hat{\rho}) \equiv \sum_m \langle m|\hat{\rho}|m\rangle \qquad \text{or} \qquad \hat{\rho}_f = \mathrm{Tr}_a(\hat{\rho}) \equiv \sum_n \langle n|\hat{\rho}|n\rangle$$

where Tr_f and Tr_a means to trace over the field modes or the atomic states respectively.

The interaction representation of the direct-product density operator is defined analogously to that for the atom or EM density operators. The unitary operator is

$$\hat{u} = \exp\left\{\frac{\hat{H}_o t}{i\hbar}\right\} \tag{5.3.7}$$

where the Hamiltonian $\hat{H}_o$ does not contain the matter-light interaction Hamiltonian

$$\hat{H}_o = \hat{H}_a + \hat{H}_f \tag{5.3.8}$$

Because $\hat{H}_a$ and $\hat{H}_f$ contain dynamical variables (operators) in distinct/disjoint spaces, the two Hamiltonians commute. The Campbell-Baker-Hausdorff theorem says

$$e^{\hat{A}+\hat{B}} = e^{\hat{A}} e^{\hat{B}} e^{[\hat{A},\hat{B}]/2}$$

so long as $\left[\hat{A},\left[\hat{A},\hat{B}\right]\right] = \left[\hat{B},\left[\hat{A},\hat{B}\right]\right]$. Therefore the unitary operators can be written as

$$\hat{u} = \exp\left\{\frac{\hat{H}_o t}{i\hbar}\right\} = \exp\left\{\frac{\hat{H}_a t}{i\hbar}\right\} \exp\left\{\frac{\hat{H}_f t}{i\hbar}\right\} = \hat{u}_a \hat{u}_f \tag{5.3.9}$$

where $\hat{u}_a$ operates only in atom-space and $\hat{u}_f$ operates only in light-space. As a note, $\hat{u}_a$ and $\hat{u}_f$ commute and can be arranged as either $\hat{u}_a \hat{u}_f$ or $\hat{u}_f \hat{u}_a$.

The interaction representation of the density operator can be written as

$$\breve{\rho} = \hat{u}^+ \hat{\rho}\, \hat{u} = \hat{u}_a^+ \hat{u}_f^+ \hat{\rho}\, \hat{u}_a \hat{u}_f \tag{5.3.10}$$

Under special circumstances when $\hat{\rho} = \hat{\rho}_a \hat{\rho}_f$, the interaction density operator can then be written as

$$\breve{\rho} = \hat{u}^+ \hat{\rho}\, \hat{u} = \hat{u}_a^+ \hat{u}_f^+ \hat{\rho}_a \hat{\rho}_f\, \hat{u}_a \hat{u}_f = \breve{\rho}_a \breve{\rho}_f$$

Topic 5.3.4: Rate Equation for the Density Operator in the Interaction Representation

The motion of the density operator in the interaction representation is important for the master equation. In fact, the equation of motion is the master equation without the fluctuation and damping terms characteristic of an interaction between a system and reservoir. In this topic and ensuing topics the probability in the density operator is assumed to be independent of time. This section demonstrates two alternative forms for the equation of motion.

Starting with Schrodiner representation $\hat{\rho} \sim |\psi(t)\rangle\langle\psi(t)|$, and using the Schrodinger equation

$$\hat{H}|\psi\rangle = i\hbar\frac{\partial}{\partial t}|\psi\rangle \qquad \text{with} \qquad \hat{H} = \hat{H}_o + \hat{V}$$

where $\hat{V}$ is an interaction potential. Subsequent sections write $\hat{H}_o = \hat{H}_s + \hat{H}_r$ as the sum of the system and reservoir Hamiltonians. The interaction potential $\hat{V}$ is the interaction energy between the system and reservoir. The derivative of the density operator gives

$$\dot{\hat{\rho}} \sim \left(\frac{\partial}{\partial t}|\psi(t)\rangle\right)\langle\psi(t)| + |\psi(t)\rangle\left(\frac{\partial}{\partial t}|\psi(t)\rangle\right)^{+} = \frac{1}{i\hbar}\left[\hat{H},\hat{\rho}\right] \tag{5.3.11}$$

This is the equation of motion of the density operator in the Schrodinger Representation

Next, to find one form of the equation of motion in the interaction representation, differentiate the interaction representation of the density operator

$$\dot{\breve{\rho}} = \frac{\partial}{\partial t}\left(\hat{u}^{+}\hat{\rho}\hat{u}\right) \qquad \text{with} \qquad \hat{u} = \exp\left(\frac{\hat{H}_o t}{i\hbar}\right)$$

to get

$$\dot{\breve{\rho}} = -\frac{\hat{H}_o}{i\hbar}\hat{u}^{+}\hat{\rho}\hat{u} + \hat{u}^{+}\frac{\partial\hat{\rho}}{\partial t}\hat{u} + \hat{u}^{+}\hat{\rho}\hat{u}\frac{\hat{H}_o}{i\hbar} = -\frac{1}{i\hbar}\left[\hat{H}_o,\breve{\rho}\right] + \hat{u}^{+}\frac{\partial\hat{\rho}}{\partial t}\hat{u} \tag{5.3.12}$$

To arrive at this result, the fact that $\left[\hat{H}_o,\hat{u}\right] = 0$ is used.

The second form of the equation comes from (5.3.12) by rearranging terms. Factoring out the unitary operators provides

$$\dot{\breve{\rho}} = -\frac{\hat{H}_o}{i\hbar}\hat{u}^{+}\hat{\rho}\hat{u} + \hat{u}^{+}\frac{\partial\hat{\rho}}{\partial t}\hat{u} + \hat{u}^{+}\hat{\rho}\hat{u}\frac{\hat{H}_o}{i\hbar} = \hat{u}^{+}\left\{-\frac{1}{i\hbar}\left[\hat{H}_o,\hat{\rho}\right] + \frac{\partial\hat{\rho}}{\partial t}\right\}\hat{u}$$

Substituting Equation 5.3.11 gives the second form of the equation of motion

$$\dot{\breve{\rho}} = \frac{1}{i\hbar}\left[\breve{V},\breve{\rho}\right] \tag{5.3.13}$$

since $\hat{u}^{+}\left(\hat{H} - \hat{H}_o\right)\hat{u} = \hat{u}^{+}\hat{V}\hat{u} = \breve{V}$. Equation 5.3.13 clearly shows that the motion of the density operator in the interaction representation is completely controlled by the interaction Hamiltonian (also in the interaction representation).

Section 5.4: Introduction to Reservoirs for Semiconductor Lasers

A typical course in lasers uses semiclassical theory (quantized matter but classical electromagnetic fields) to determine the rate of stimulated emission from a group of atoms. The procedure finds a solution to the Liouville equation

$$\frac{\partial \rho_{mn}}{\partial t} = \frac{1}{i\hbar}\left[\hat{H}_o, \hat{\rho}\right]_{mn} - \frac{\rho_{mn} - \overline{\rho}_{mn}}{\tau_{mn}} \tag{5.4.1}$$

where $\tau_{nn} = \tau$ is the population relaxation time and $\tau_{mn} = T_2$ is the dipole dephasing time which leads to a non-zero width for the gain curve. The Hamiltonian $\hat{H}_o = \hat{H}_a + \hat{V}$ consists of the free-atom Hamiltonian $\hat{H}_a$ and the semi-classical matter-light interaction $\hat{V}$. The classical theory of EM fields does not recognize the inherent uncertainty in conjugate variables such as amplitude and phase. The classical theory obtains from the quantum theory by allowing the number of photons (or equivalently, amplitude) to be very large. However the semiclassical theory assumes that the conjugate variables commute even at low field levels, clearly contrary to the quantum theory.

Recall that the Liouville equation (5.4.1) is a result of writing the Hamiltonian as

$$\hat{H} = \hat{H}_o + \hat{H}_{other} = \hat{H}_a + \hat{V} + \hat{H}_{other}$$

which leads to

$$\frac{\partial \hat{\rho}}{\partial t} = \frac{1}{i\hbar}\left[\hat{H}_o, \hat{\rho}\right] + \left(\frac{\partial \hat{\rho}}{\partial t}\right)_{other} = \frac{1}{i\hbar}\left[\hat{H}_o, \hat{\rho}\right] + \left(\frac{\partial \hat{\rho}}{\partial t}\right)_{pump} + \left(\frac{\partial \hat{\rho}}{\partial t}\right)_{coll} + \left(\frac{\partial \hat{\rho}}{\partial t}\right)_{spont}$$

The density operator provides both a macroscopic statistical average and the microscopic quantum mechanical average. The $\hat{H}_{other}$ term indicates that an "external" agent acts on the system. These external agents cause the relaxation effects. The external agents can be modeled as "reservoirs" (similar concept to thermal reservoirs). Subsequent sections show that a reservoir induces rapid fluctuations (Langevin noise) as well as damping. This section introduces the notion of a reservoir and discusses the associated fluctuation-dissipation theorem.

Topic 5.4.1: Definition of Reservoir

The complete system is divided into a small system under study and a collection of reservoirs. The reservoirs are large systems that provide equilibrium for the smaller system. A reservoir is a system with an extremely large number of degrees of freedom. For example, a reservoir of two-level atoms or harmonic oscillators necessarily contains a large number of atoms or oscillators. A reservoir of light consists of a set of modes where the number of such modes is extremely large. Typically, a specific energy distribution is assumed to exist in the reservoir. For example, if the reservoir consists of point particles (such as gas molecules) then one might assume a Boltzmann distribution for the energy.

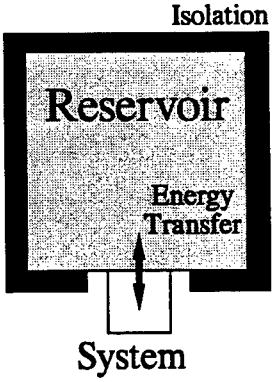


Figure 5.4.1: The reservoir can exchange energy with the system under study.

The reservoir is brought into contact with the small system so that energy can flow between the system and the

reservoir. The reservoir has such a large number of degrees-of-freedom that any energy transferred from the small system has negligible affect on the reservoir energy distribution. For a concrete example, suppose the small system consists of a single gas molecule and the reservoir has a large number of molecules all at thermal equilibrium (i.e., a Boltzmann energy distribution). The temperature of the small system will eventually match the temperature of the larger system. However, temperature is a measure of kinetic energy. Therefore, to say that the temperatures are the same is to say that the average kinetic energy of the single molecule is the same the average kinetic energy of all the molecules in the reservoir (some move faster than others).

Suppose the molecule in the small system has a much larger than average kinetic energy (maybe a factor of 10). The extra energy is eventually transferred to the reservoir. This extra energy is distributed to all of the molecules in the reservoir, which makes negligible changes in the total reservoir distribution. Essentially the initial packet of transferred energy is lost from the small system—never to return—because it is distributed over so many degrees of freedom in the large system. In effect, the reservoir has "absorbed" the "extra" system energy and the motion of single molecule is "damped".

The reservoir energy distribution defines average quantities for the reservoir. The contact between the two systems brings the small system into equilibrium with the reservoir, which therefore defines the average quantities for the small system. Suppose the single atom in the small system is initially in equilibrium with the reservoir. Occasionally, a large chunk of energy will be transferred from the reservoir to the small system as a thermal fluctuation. As a result, the single atom will have more energy than its equilibrium value. Eventually this extra energy will be damped out. The correlation between fluctuations is assumed to occur on such short times scales as to be negligible. The process of transferring energy between the small system and the reservoir is an example application of the fluctuation-dissipation theorem. The theorem basically states that a reservoir (or other system) both damps the small system and induces fluctuations in the small system. The two processes go together and cannot be separated. Often, on a phenomenological level, the fluctuations are included in rate equations through a Langevin function.

Topic 5.4.2: An example for the Fluctuation-Dissipation Theorem

For example, Brownian motion of a small particle in a liquid consists of rapid, uncorrelated movements. The motion of the small particle is a result of the interaction between the particle and its liquid environment, which acts as a reservoir. The phenomenological equations are derived from Newton's second law. For one-dimensional motion

$$m\ddot{x} = -m\gamma\dot{x} + f(t) \qquad (5.4.2)$$

where the term proportional to the velocity is the damping term. The damping is intimately related to the Langevin force f(t) even though the above phenomenological equation suggests that they are separate issues. When one term is present, the other one must likewise be there. The Langevin force f(t) is rapidly varying with the average value of $\langle f \rangle = 0$.

The fluctuations associated with the Langevin force are "stationary" with exceedingly small correlation times. Stationary means that the probability distribution P for the fluctuations does not change with time. An ergodic process assumes that the average of a function f(t) can be computed by either

$$\langle y(t)\rangle = \frac{1}{\tau}\int_0^{\tau} dt\ y(t) \qquad \text{(for sufficiently large } \tau\text{)}$$

or by using the notion of ensemble

$$\langle y\rangle = \int dy\ y\ P(y)$$

where P(y) is the probability density. The average of Equation 5.4.2 can be handled by either method. To use the time average, the time interval τ is assumed long compared with the correlation time but short compared with the time scale of interest. In this way, $\langle f\rangle = 0$ but $\langle x\rangle$ can still depend on time τ.

Taking the average of Equation 5.4.2 gives

$$m\frac{d^2}{dt^2}\langle x\rangle + m\gamma\frac{d}{dt}\langle x\rangle = \langle f(t)\rangle = 0$$

or equivalently

$$m\frac{d}{dt}v + m\gamma v = 0 \qquad \text{where} \qquad v = \frac{d}{dt}\langle x\rangle$$

Therefore, the simple differential equation has the solution

$$v(t) = v(0)\exp(-\gamma t)$$

Likewise, it is possible to show that the variance of the motion is not zero even though the initial macroscopic velocity is zero. The variance is nonzero because of the fluctuations induced in the velocity of the particle by the reservoir.

Topic 5.4.3: Reservoirs for a Laser

What are the reservoirs for the laser? First consider collisions. For simplicity, consider a laser with a single atom. Assume the single atom is embedded in a "background material" such as a crystal. The reservoir might consists of phonons that exist on the lattice or free electrons that can participate in collisions. The "background material" as part of the laser has a specified temperature and composition. This means that the reservoir can be assumed to have a Boltzmann distribution characterized by a certain temperature T.

The reservoir for spontaneous emission is the collection of all optical modes in space—an extremely large number. The mode is though of as a place to dump photons. A mode is characterized by a given wave vector and polarization. They are typically pictured as "empty" travelling waves (the quantum vacuum); that is, there are no photons in the mode. When an excited single atom emits a photon, the photon is "absorbed" by the reservoir and never allowed to interact with the atom again. For this reason, people refer to the interaction of the (spontaneous emission) reservoir and atom as an irreversible process. The spontaneous emission process can be reversed only if the emitted photon can interact with the atom again. The re-absorption of the photon by the atom can be accomplished if the atom is in a Fabry-Perot cavity, for example.

Topic 5.4.4: Comment

The Liouville equation for the density operator is essentially a differential equation for the energy level occupation number (i.e., $\langle n|\hat{\rho}|n\rangle$) and the induced polarization (off-diagonal terms). The following sections show how a quantum mechanical reservoir gives rise to damping and fluctuation terms in the Liouville equation (a.k.a., master equation). The damping term appears as

$$\left.\frac{\partial\hat{\rho}}{\partial t}\right|_{other}$$

while, as expected, the average of the fluctuations disappears. The next sections will use the trace over the reservoir states to calculate the average. The tracing operation produces a zero average for the fluctuations and removes the reservoir degrees of freedom from the differential equation.

The formalism can be applied to spontaneous and stimulated emission from atoms. The density operator is used so that various types of EM states (Fock, coherent and squeezed) can be used as well as various atomic states. The density operator/matrix accounts for all possible knowledge of the system.

Section 5.5: Quantum Mechanical Fluctuation-Dissipation Theorem

The fluctuation-dissipation theorem incorporates the role of the reservoir in the master equation. The influence of the reservoirs become apparent for the applications of the density operator. The laser rate equations provide one example where relaxation terms describe the decay of the carrier and lead to a cavity lifetime. Haken (and others) clearly show how the dissipation-fluctuations arise from quantum mechanical interactions between a system and reservoir. In particular, the Langevin source terms appear as the Fourier transforms of the reservoir-system coupling coefficients. For the Langevin forces to be truly delta-function correlated, the coupling strengths (as a function of frequency) must be broadband (approximately independent of frequency). The jitter and relaxation effects are most clearly seen in the Hesienberg representation of the raising/lowering operators; this is probably because of the close relation between operators in the Heisenberg picture and the classical counterparts.

Consider a small system with a single harmonic oscillator and a reservoir with a very large number of harmonic oscillators. Let $\left\{\hat{R}_\omega^+\right\}$ and $\left\{\hat{R}_\omega^-\right\}$ be the creation and annihilation operators for mode ω in the reservoir (in the Schrodinger representation). Let $\hat{S}^+$ and $\hat{S}^-$ be the *raising* and *lowering* operators for the harmonic oscillator in the small system. The Hamiltonian for the system and reservoir $\hat{H} = \hat{H}_s + \hat{H}_r + \hat{H}_{sr}$ consist of the Hamiltonians for the system, reservoir and system-reservoir interaction

$$\hat{H}_s = \hbar\omega_s\left(\hat{S}^+\hat{S}^- + \frac{1}{2}\right) \qquad \hat{H}_r = \sum_\omega \hbar\omega\left(\hat{R}_\omega^+\hat{R}_\omega^- + \frac{1}{2}\right)$$

$$\hat{H}_{sr} = \sum_\omega \beta_\omega \hat{S}^+\hat{R}_\omega^- + \sum_\omega \beta_\omega^*\hat{S}^-\hat{R}_\omega^+ = \hat{S}^+\sum_\omega \beta_\omega\hat{R}_\omega^- + \hat{S}^-\sum_\omega \beta_\omega^*\hat{R}_\omega^+$$

The system-reservoir interaction Hamiltonian consists of terms that explicitly conserve energy. For example, the annihilation operator $\left\{\hat{R}_\omega^-\right\}$ removes a quantum of energy from the reservoir and the raising operator $\hat{S}^+$ increase the energy of the system. The β_ω are the coupling strength of a reservoir oscillator with angular frequency ω to the system. The operators obey the following commutation relations.

$$\left[\hat{S}^-,\hat{S}^-\right] = 0 = \left[\hat{S}^+,\hat{S}^+\right] \qquad \left[\hat{S}^-,\hat{S}^+\right] = 1$$

$$\left[\hat{R}_i^-,\hat{R}_j^-\right] = 0 = \left[\hat{R}_i^+,\hat{R}_j^+\right] \qquad \left[\hat{R}_i^-,\hat{R}_j^+\right] = \delta_{ij}$$

which also apply for equal-time commutators in the Heisenberg representation. The Heisenberg representation of an operator $\hat{O}$ is given by

$$\breve{O} = \hat{u}^+\hat{O}\hat{u}$$

where

$$\hat{u} = \exp\left(\frac{\hat{H}t}{i\hbar}\right)$$

Note the use of the symbol over the operator "$\breve{O}$" to indicate the Heisenberg representation. The total Hamiltonian in the Schrodinger and Heisenberg representations is the same since $\hat{u}$ and $\hat{H}$ commute

$$\breve{H} = \hat{u}^+\hat{H}\hat{u} = \hat{u}^+\hat{u}\hat{H} = \hat{H}$$

However, this is not necessarily true for each individual Hamiltonian such as $\hat{H}_{sr}$. The Heisenberg Hamiltonians are obtained just by adding the "inverted carat" to each operator. The equation of motion for a Heisenberg operator $\breve{O}$ is

$$\frac{d\breve{O}}{dt} = \frac{i}{\hbar}\left[\breve{H},\breve{O}\right] + \hat{u}^{+}\frac{\partial \hat{O}}{\partial t}\hat{u}$$

Usually the last term is zero for most operators (except possibly the density operator).

The operators of the theory therefore satisfy the following rate equations

$$\dot{\breve{S}}^{-} = \frac{i}{\hbar}\left[\breve{H},\breve{S}^{-}\right] = \frac{i}{\hbar}\left[\breve{H}_{s} + \breve{H}_{r} + \breve{H}_{sr},\breve{S}^{-}\right] = \frac{i}{\hbar}\left[\breve{H}_{s} + \breve{H}_{sr},\breve{S}^{-}\right] \qquad (5.5.1)$$

The first commutator is

$$\left[\breve{H}_{s},\breve{S}^{-}\right] = \left[\hbar\omega_{s}\left(\breve{S}^{+}\breve{S}^{-} + \frac{1}{2}\right),\breve{S}^{-}\right] = \hbar\omega_{s}\left[\breve{S}^{+}\breve{S}^{-},\breve{S}^{-}\right] = \hbar\omega_{s}\left[\breve{S}^{+},\breve{S}^{-}\right]\breve{S}^{-} = -\hbar\omega_{s}\breve{S}^{-}$$

The second commutator is

$$\left[\breve{H}_{sr},\breve{S}^{-}\right] = \left[\breve{S}^{+}\sum_{\omega}\beta_{\omega}\breve{R}_{\omega}^{-} + \breve{S}^{-}\sum_{\omega}\beta_{\omega}^{*}\breve{R}_{\omega}^{+},\breve{S}^{-}\right] = \left[\breve{S}^{+}\sum_{\omega}\beta_{\omega}\breve{R}_{\omega}^{-},\breve{S}^{-}\right] = \left[\breve{S}^{+},\breve{S}^{-}\right]\sum_{\omega}\beta_{\omega}\breve{R}_{\omega}^{-} = -\sum_{\omega}\beta_{\omega}\breve{R}_{\omega}^{-}$$

Combining the commutators into Equation 5.5.1 yields

$$\dot{\breve{S}}^{-} = -i\omega_{s}\breve{S}^{-} - \frac{i}{\hbar}\sum_{\omega}\beta_{\omega}\breve{R}_{\omega}^{-} \qquad (5.5.2)$$

The adjoint of Equation 5.5.2 provides a similar equation for the raising operator.

$$\dot{\breve{S}}^{+} = i\omega_{s}\breve{S}^{+} + \frac{i}{\hbar}\sum_{\omega}\beta_{\omega}^{*}\breve{R}_{\omega}^{+} \qquad (5.5.3)$$

It is also necessary to find the equations of motion for the reservoir operators.

$$\dot{\breve{R}}_{\omega}^{-} = \frac{i}{\hbar}\left[\breve{H},\breve{R}_{\omega}^{-}\right] = \frac{i}{\hbar}\left[\breve{H}_{s} + \breve{H}_{r} + \breve{H}_{sr},\breve{R}_{\omega}^{-}\right] = \frac{i}{\hbar}\left[\breve{H}_{r} + \breve{H}_{sr},\breve{R}_{\omega}^{-}\right] \qquad (5.5.4)$$

The first commutator provides

$$\left[\breve{H}_{r},\breve{R}_{\omega}^{-}\right] = \left[\sum_{\omega'}\hbar\omega'\left(\breve{R}_{\omega'}^{+}\breve{R}_{\omega'}^{-} + \frac{1}{2}\right),\breve{R}_{\omega}^{-}\right] = \sum_{\omega'}\hbar\omega'\left[\breve{R}_{\omega'}^{+},\breve{R}_{\omega}^{-}\right]\breve{R}_{\omega'}^{-} = \sum_{\omega'} -\hbar\omega'\delta_{\omega\omega'}\breve{R}_{\omega}^{-} = -\hbar\omega\breve{R}_{\omega}^{-}$$

The second commutator gives

$$\left[\breve{H}_{sr},\breve{R}_{\omega}^{-}\right] = \left[\breve{S}^{+}\sum_{\omega'}\beta_{\omega'}\breve{R}_{\omega'}^{-} + \breve{S}^{-}\sum_{\omega'}\beta_{\omega'}^{*}\breve{R}_{\omega'}^{+},\breve{R}_{\omega}^{-}\right] = \left[\breve{S}^{-}\sum_{\omega'}\beta_{\omega'}^{*}\breve{R}_{\omega'}^{+},\breve{R}_{\omega}^{-}\right] = -\beta_{\omega}^{*}\breve{S}^{-}$$

Combining the last two commutators into Equation 5.5.4 gives

$$\dot{\breve{R}}_{\omega}^{-} = -i\omega\breve{R}_{\omega}^{-} - \frac{i}{\hbar}\beta_{\omega}^{*}\breve{S}^{-} \qquad (5.5.5)$$

The adjoint gives

$$\dot{\breve{R}}_{\omega}^{+} = i\omega\breve{R}_{\omega}^{+} + \frac{i}{\hbar}\beta_{\omega}\breve{S}^{+} \qquad (5.5.6)$$

An equation for the lowering operator (for the system) can be found by combining Equations 5.5.5 and 5.5.2. First, formally solve equation 5.5.5 by using an integrating factor. As reviewed in Appendix 1, a first order differential equation $\dot{y} - ay = f(t)$ has the solution

$$y(t) = \frac{\mu(0)y(0)}{\mu(t)} + \frac{1}{\mu(t)}\int_0^t d\tau\, \mu(\tau) f(\tau)$$

with $\mu(t) = e^{-at}$. Equation 5.5.5 has the integrating factor $\mu = e^{i\omega t}$ and the formal solution

$$\breve{R}_\omega^-(t) = e^{-i\omega t}\breve{R}_\omega^-(0) + \frac{i\beta_\omega^*}{\hbar}\int_0^t d\tau\, e^{-i\omega(t-\tau)}\breve{S}^-(\tau) \qquad (5.5.7)$$

Substituting Equation 5.5.7 into Equation 5.5.2 provides

$$\dot{\breve{S}}^- + i\omega_s\breve{S}^- + \frac{1}{\hbar^2}\sum_\omega |\beta_\omega|^2 \int_0^t d\tau\, e^{-i\omega(t-\tau)}\breve{S}^-(\tau) = -\frac{i}{\hbar}\sum_\omega \beta_\omega e^{-i\omega t}\breve{R}_\omega^-(0)$$

which can be rewritten as

$$\dot{\breve{S}}^- + i\omega_s\breve{S}^- + \frac{1}{\hbar^2}\int_0^t d\tau\, \breve{S}^-(\tau)\sum_\omega |\beta_\omega|^2 e^{-i\omega(t-\tau)} = -\frac{i}{\hbar}\sum_\omega \beta_\omega e^{-i\omega t}\breve{R}_\omega^-(0) \qquad (5.5.8)$$

The integral in Equation 5.5.8 gives the damping whereas the summation is the Langevin fluctuation. First evaluate the integral. The summation in the integrand can be evaluated. Let $g(\omega)$ be the density of states (i.e., the number of β_ω per unit frequency range).

$$\sum_\omega |\beta_\omega|^2 e^{-i\omega(t-\tau)} = \int_x^y d\omega\, g(\omega)|\beta_\omega|^2 e^{-i\omega(t-\tau)} \qquad (5.5.9)$$

where x,y are the smallest and largest allowed frequencies, respectively. A typical assumption is that $g(\omega)|\beta_\omega|^2$ is essentially independent of frequency, which is usually stated as the coupling strength β_ω being relatively constant. For a number of calculations, the upper limit on the integral remains finite, which avoids pesky infinities. A good example is for the Casimir effect (refer to Milonni's book on the Quantum Vacuum). We take the upper limit to be infinity for simplicity. The summation becomes

$$\sum_\omega |\beta_\omega|^2 e^{-i\omega(t-\tau)} = \int_x^\infty d\omega\, g(\omega)|\beta_\omega|^2 e^{-i\omega(t-\tau)} = g|\beta|^2 \int_x^\infty d\omega\, e^{-i\omega(t-\tau)} \qquad (5.5.10)$$

Several texts and papers handle the integral differently. Some authors include negative frequencies in the integral to obtain a Dirac delta function; however, it is non-physical to have negative frequencies since this corresponds to negative energies for the system harmonic oscillator.

$$\int_x^\infty d\omega\, e^{-i\omega(t-\tau)} = \begin{cases} 2\pi\delta(t-\tau) & x = -\infty \\ \pi g|\beta|^2\delta(t-\tau) - ig|\beta|^2\dfrac{P}{t-\tau} & x = 0 \end{cases} \qquad (5.5.11)$$

The "P" in the above relation refers to the principal part. The development in this book uses the non-physical expression with the negative frequencies.

We obtain the fluctuation-damping expression for the system operators by substituting Equations 10.43.3-5 into Equation 10.32.2.

$$\dot{\breve{S}}^- + i\omega_s\breve{S}^- + \frac{g|\beta|^2 2\pi}{\hbar^2}\int_0^t d\tau\, \breve{S}^-(\tau)\delta(t-\tau) = -\frac{i}{\hbar}\sum_\omega \beta_\omega e^{-i\omega t}\breve{R}_\omega^-(0) \qquad (5.5.12)$$

The definition of the Dirac delta function

$$\int_0^b d\tau\,\delta(\tau-a)=\begin{cases}1 & a\in(0,b)\\ 1/2 & a=b,0\\ & a\notin[a,b]\end{cases}$$

the equation of motion becomes

$$\dot{\breve{S}}^- + i\omega_s\breve{S}^- + \frac{\pi g|\beta|^2}{\hbar^2}\breve{S}^-(t) = -\frac{i}{\hbar}\sum_\omega \beta_\omega e^{-i\omega t}\breve{R}_\omega^-(0) \qquad (5.5.13)$$

The frequency of oscillation is controlled by the second term which is $i\omega_s\breve{S}^-$. The damping/relaxation is controlled by the third term

$$\frac{\pi g|\beta|^2}{\hbar^2}\breve{S}^-(t)$$

The damping coefficient is

$$\gamma = \frac{\pi g|\beta|^2}{\hbar^2}$$

The adjoint of Equation 5.5.12 provides the equation of motion for the raising operator.

$$\dot{\breve{S}}^+ - i\omega_s\breve{S}^+ + \frac{\pi g|\beta|^2}{\hbar^2}\breve{S}^+(t) = \frac{i}{\hbar}\sum_\omega \beta_\omega^* e^{+i\omega t}\breve{R}_\omega^+(0) \qquad (5.5.14)$$

The last term in Equation 5.5.13 is the Langevin "force" describing the random fluctuations that the reservoir induces in the system. The Langevin force term

$$\Gamma(t) = -\frac{i}{\hbar}\sum_\omega \beta_\omega e^{-i\omega t}\breve{R}_\omega^-(0)$$

averages to zero. The average can be calculated by either taking the time average or the ensemble average. For ergodic systems, the averages are the same. The fluctuations induced by the reservoir are stationary and ergodic. A time average must be over a sufficiently long time that any correlation in the fluctuations of the reservoir has died out. An "ensemble of reservoirs" has reservoirs in every possible internal configuration. This means that there is a reservoir in the system with any particular value of Γ. The ensemble average is

$$\langle\Gamma(t)\rangle_{reserv} = Tr(\rho_r\Gamma(t))$$

The density operator $\hat{\rho}_r$ is defined through a Boltzmann distribution

$$\hat{\rho}_r = \frac{1}{Z}\exp\left(-\frac{\hat{H}_r}{k_B T}\right)$$

where Z is the normalization

$$Z = Tr_r\left\{\exp\left(-\frac{\hat{H}_r}{k_B T}\right)\right\}$$

Appendix 6 shows how the reservoir density operator is related to the Boltzmann distribution. Actually, for this calculation, we really only require that the density operator be diagonal (a pure state). For a reservoir consisting of a very large number of Harmonic oscillators with many possible frequencies, the Hamiltonian is

$$\hat{H}_r = \sum_\omega \hbar\omega\left(\hat{R}_\omega^+\hat{R}_\omega^- + \frac{1}{2}\right)$$

and a Fock state is denoted by

$$|\{n\}\rangle = |n_1, n_2, n_3, \ldots\rangle$$

The average of the fluctuation term becomes (similar to Appendix 6)

$$\langle\Gamma(t)\rangle_{reserv} = \mathrm{Tr}(\rho_r\Gamma(t)) = -\frac{i}{\hbar}\mathrm{Tr}\left(\rho_r\sum_\omega\beta_\omega e^{-i\omega t}\breve{R}_\omega^-(0)\right) = -\frac{ie^{-i\omega t}}{\hbar Z}\sum_\omega\beta_\omega\,\mathrm{Tr}_r\left\{\exp\left(-\frac{\hat{H}_r}{k_BT}\right)\hat{R}_\omega^-\right\}$$

Notice the notation of $\hat{R}_\omega^- = \breve{R}_\omega^-(0)$ is used. The trace is easy to calculate to find

$$\mathrm{Tr}_r\left\{\exp\left(-\frac{\hat{H}_r}{k_BT}\right)\hat{R}_\omega^-\right\} = \sum_{\{n\}}\langle\{n\}|\exp\left(-\frac{\hat{H}_r}{k_BT}\right)\hat{R}_\omega^-|\{n\}\rangle$$

Inserting the closure relation provides

$$\mathrm{Tr}_r\left\{\exp\left(-\frac{\hat{H}_r}{k_BT}\right)\hat{R}_\omega^-\right\} = \sum_{\{n\},\{m\}}\langle\{n\}|\exp\left(-\frac{\hat{H}_r}{k_BT}\right)|\{m\}\rangle\langle\{m\}|\hat{R}_\omega^-|\{n\}\rangle$$

Here's where the fact that $\hat{\rho}_r$ is diagonal becomes important.

$$\mathrm{Tr}_r\left\{\exp\left(-\frac{\hat{H}_r}{k_BT}\right)\hat{R}_\omega^-\right\} = \sum_{\{n\},\{m\}}\exp\left(-\frac{1}{k_BT}\sum_{\{m\}}E_{\{m\}}\right)\langle\{n\}|\{m\}\rangle\langle\{m\}|\hat{R}_\omega^-|\{n\}\rangle$$

where $E_{\{m\}}$ is an obvious notation. The orthonormality of the Fock states provides

$$\langle\{n\}|\{m\}\rangle = \delta_{\{n\},\{m\}}$$

The trace becomes

$$\mathrm{Tr}_r\left\{\exp\left(-\frac{\hat{H}_r}{k_BT}\right)\hat{R}_\omega^-\right\} = \sum_{\{n\},\{m\}}\exp\left(-\frac{1}{k_BT}\sum_{\{m\}}E_{\{m\}}\right)\langle\{n\}|\hat{R}_\omega^-|\{n\}\rangle$$

Now its clear that the trace operation gives zero because the annihilation operator removes a quanta of energy from the ket . For example

$$\langle n_1, n_2\ldots|\hat{R}_1^-|n_1, n_2\ldots\rangle = \langle n_1, n_2\ldots|n_1 - 1, n_2\ldots\rangle = 0$$

since Fock states with unequal occupation numbers are orthogonal. Therefore the expectation value of the Langevin fluctuation is

$$\langle\Gamma(t)\rangle_{reserv} = \mathrm{Tr}(\rho_r\Gamma(t)) = 0$$

At this point we can illustrate how the correlation function plays a pivotal role for the population number (population inversion) and the polarization (gain).

Section 5.6: The Master Equation

This section shows how a group of reservoirs can be used to determine the steady-state for a quantum mechanical system. It shows how the damping and fluctuation terms arise in the rate equation for the density operator. There are three approaches to the fully quantum mechanical rate equations. (1) The Heisenberg equations for the creation and annihilation operators is perhaps the most easily compared with the classical rate equation with the damping and Langevin terms. (2) The density operator approach is most appropriate for applying quantum mechanical fields to a collection of atoms. The density operator is adjusted to the type of applied field. (3) A distribution approach. The previous section discusses the results for the Heisenberg representation and this one discusses the density operator approach.

Topic 5.6.1: Brief Plan of Approach

The complete system is divided into a small system and a number of reservoirs. The discussion in this section assumes a two-level atom capable of interacting with an EM field and a set of reservoirs. The system consists of the two-level atom and the interacting optical field. The Hamiltonian for the system and reservoirs can be written as

$$H = H_{atom} + H_{light} + H_{\substack{atom\\lite}} + H_{\substack{atom\\reserv}} + H_{\substack{light\\reserv}} + H_{\substack{mult\\reserv}} = H_s + H_{sr} + H_r$$

where $H_s = H_{atom} + H_{light} + H_{\substack{atom\\light}}$, $H_{sr} = H_{\substack{atom\\reserv}} + H_{\substack{light\\reserv}}$. The various Hamiltonians H_{atom}, H_{light}, $H_{\substack{atom\\light}}$, H_{sr}, $H_{\substack{atom\\reserv}}$, $H_{\substack{light\\reserv}}$, $H_{\substack{mult\\reserv}}$ are for the free atom (i.e., atom by itself), free light (i.e., a light field by itself), atom-light interaction (stimulated or spontaneous emission), atom-reservoir interaction, light-reservoir interaction, and the "free" Hamiltonian for the group of reservoirs, respectively. The atom-light interaction is initiated at time t_o. The density operator exists in a direct product space. Up to and including this time t_o, the density operator can be written as $\rho = \rho_{atom}\rho_{light}\rho_{res1}\rho_{res2}\cdots$.

The rate equation for the density operator using the total Hamiltonian can be written as

$$\dot{\rho} = \frac{1}{i\hbar}\left[\hat{H},\rho\right] = \frac{1}{i\hbar}\left[\hat{H}_s,\rho\right] + \frac{1}{i\hbar}\left[\hat{H}_{sr},\rho\right] + \frac{1}{i\hbar}\left[\hat{H}_r,\rho\right] \qquad (5.6.1)$$

where H_s does not involve any of the reservoir operators. Taking the trace of Equation 5.6.1 over the reservoir states then reduces $\dot{\rho}$ and $\left[\hat{H},\rho\right]$ to terms involving only the *system* density operator ρ_s. The last two terms, previously called "other", provide the "relaxation" effects. Equation 5.6.1 becomes

$$\dot{\rho}_s = \frac{1}{i\hbar}\left[\hat{H}_s,\rho_s\right] + \left.\frac{\partial \rho_s}{\partial t}\right|_{other} \qquad (5.6.2)$$

where

$$\left.\frac{\partial \rho_s}{\partial t}\right|_{other} = \sum_i \left.\frac{\partial \rho_s}{\partial t}\right|_{\substack{reserv\\ \#i}}$$

Equation 5.6.2 is similar to the Liouville equation discussed in the first two sections of this chapter.

Topic 5.6.2: Multiple Reservoirs

Consider the multiple reservoirs interacting with a single system. Let the operators corresponding to the a^{th} reservoir be $R^{(a)}=\{R_i^{(a)}\}$. Each reservoir can have a different temperature. Segregate the operators for the system so that $S^{(a)}=\{S_i^{(a)}\}$ are linked with the a^{th} reservoir; i.e., the degrees of freedom represented by set $S^{(a)}$ interact with the a^{th} reservoir. All of the reservoir operators commute with the system operators for all times in the Schrodinger and interaction representations.

$$\left[\hat{S}_i^{(a)},\hat{R}_i^{(b)}\right]=0$$

However, by necessity, each individual set contains non-commuting operators. For example, there are indices "i" and "j" such that

$$\left[\hat{S}_i^{(a)},\hat{S}_j^{(b)}\right]\neq 0$$

and something similar for the set R. Assume that there is a total of "N" reservoirs. The interaction Hamiltonian is assumed to be

$$\hat{H}_{sr}=\hat{V}=\sum_a\sum_i S_i^{(a)}R_i^{(a)}$$

Notice that in this topic there is no distinction made between raising and lowering operators.

Example: Suppose the system consists of an atom and the spontaneous emission electromagnetic field, which interacts with the vacuum (as a reservoir #1). The "free" field Hamiltonian is

$$\hat{H}=\sum_k \hbar\omega_k\left(\hat{b}_k^+\hat{b}_k^-+\frac{1}{2}\right)$$

There are two operators $\hat{S}_i^{(1)}$ and $\hat{S}_j^{(1)}$ in the set "S" such that, for each k,

$$\hat{S}_i^{(1)}=\hat{b}_k^+ \quad \text{and} \quad \hat{S}_j^{(2)}=\hat{b}_k^-$$

Assume that the system and reservoir are not correlated at t=0 when the interaction begins. The density operator at t=0 can be written as

$$\hat{\rho}(0)=\hat{\rho}^{(s)}(0)\hat{\rho}^{(1)}(0)\hat{\rho}^{(2)}(0)\ldots\hat{\rho}^{(N)}(0)=\hat{\rho}^{(s)}(0)\bigotimes_{r=1}^{N}\hat{\rho}^{(r)}(0)\equiv\hat{\rho}^{(s)}(0)\,\hat{\rho}^{(1\ldots N)}(0)$$

where $\rho^{(s)}(0)$ is the system density operator and $\hat{\rho}^{(1\ldots N)}(0)$ is defined to be the density operator of the "N" reservoirs at t=0. For all times (not just t=0) it is permissible to write

$$\hat{\rho}^{(s)}(t)=\mathrm{Tr}_1\mathrm{Tr}_2\cdots\mathrm{Tr}_N\hat{\rho}(t)\equiv\mathrm{Tr}_{1\ldots N}\hat{\rho}(t)$$

Topic 5.6.3: Dynamics and the Perturbation Expansion

The total density operator $\hat{\rho}$ in the Schrodinger representation satisfies

$$\dot{\hat{\rho}} = \frac{1}{i\hbar}\left[\hat{H},\hat{\rho}\right] \tag{5.6.3}$$

as discussed in Topic 5.3.4. In this section, symbols such as $\hat{O}$ are operators in Schrodinger representation, and $\breve{O}$ are operators in the interaction representation. It would be nice to use a perturbation expansion to find the density operator in Equation 5.6.3 by treating the right-hand side as small. Working in the interaction representation facilitates the procedure. Equation 5.3.13 in Topic 5.3.4 says the motion of $\breve{\rho}$ in Hilbert space is controlled by the interaction Hamiltonian in the interaction representation according to

$$\dot{\breve{\rho}} = \frac{1}{i\hbar}\left[\breve{H}_{sr},\breve{\rho}\right] \tag{5.6.4}$$

The interaction representation quantities are

$$\breve{\rho} = \hat{u}^{+}\hat{\rho}\hat{u} \qquad\qquad \breve{H}_{sr} = \hat{u}^{+}\hat{H}_{sr}\hat{u}$$

where the explicit form of the unitary operator $\hat{u} = \exp\left(\frac{\hat{H}_o t}{i\hbar}\right)$ assumes the reservoir-system interaction starts at $t_o = 0$. The "free" Hamiltonian is $\hat{H}_o = \hat{H}_s + \hat{H}_r$, which consists of the evolution of the system and reservoirs without any interaction between them. The evolution operator can be written as $\hat{u} = \hat{u}_s\hat{u}_r$ since the two Hamiltonians commute $\left[\hat{H}_s,\hat{H}_r\right] = 0$ and the exponential can therefore be divided into two pieces. Equation 5.6.4 can be integrated to give

$$\breve{\rho}(t) = \breve{\rho}(0) + \frac{1}{i\hbar}\int_0^t d\tau\left[\breve{H}_{sr}(\tau),\breve{\rho}(\tau)\right] \tag{5.6.5}$$

Assume that $\breve{H}_{sr}$ is small. Equation 5.6.5 can be repeatedly substituted into Equation 5.6.4 to get

$$\dot{\breve{\rho}}(t) = \frac{1}{i\hbar}\left[\breve{H}_{sr}(t),\breve{\rho}(0)\right] + \frac{1}{(i\hbar)^2}\int_0^t d\tau\left[\breve{H}_{sr}(t),\left[\breve{H}_{sr}(\tau),\breve{\rho}(0)\right]\right] + \ldots \tag{5.6.6}$$

Again it is apparent that the interaction Hamiltonian is the sole motivator for the evolution of the density operator in the interaction representation. For convenience of notation, substitute $V = H_{sr}$

$$\dot{\breve{\rho}} \cong \frac{1}{i\hbar}\left[\breve{V},\breve{\rho}(0)\right] + \frac{1}{(i\hbar)^2}\int_0^t dt'\left[\breve{V}(t),\left[\breve{V}(t'),\breve{\rho}(0)\right]\right] + \ldots \tag{5.6.7}$$

It is necessary to take the trace of Equation 5.6.7 over the reservoir states to find the equation of motion for the system density operator. There are two ways to accomplish the task. First consider that the eigenvector are always independent of time. If $\{|n\rangle\}$ are the reservoir eigenstates then

$$\mathrm{Tr}\left(\dot{\breve{\rho}}\right)=\sum_{n}\langle n|\dot{\breve{\rho}}|n\rangle=\frac{\partial}{\partial t}\sum_{n}\langle n|\breve{\rho}|n\rangle=\frac{\partial}{\partial t}\mathrm{Tr}\;\breve{\rho}=\dot{\breve{\rho}}_s$$

As a second method, the same expression is easily seen to hold by using the definition of derivative

$$\mathrm{Tr}_r\;\dot{\breve{\rho}}(t)\cong\mathrm{Tr}_r\;\frac{\breve{\rho}(t+\Delta t)-\breve{\rho}(t)}{\Delta t}=\frac{1}{\Delta t}\left\{\mathrm{Tr}_r\;\breve{\rho}(t+\Delta t)-\mathrm{Tr}_r\;\breve{\rho}(t)\right\}=\frac{1}{\Delta t}\left\{\breve{\rho}_s(t+\Delta t)-\breve{\rho}_s(t)\right\}\cong\dot{\breve{\rho}}_s(t)$$

Either way, taking the trace over all of the reservoirs gives (refer to the next paragraph)

$$\dot{\breve{\rho}}_s=\frac{1}{i\hbar}\mathrm{Tr}_{1\ldots N}\left[\breve{V},\breve{\rho}(0)\right]+\frac{1}{(i\hbar)^2}\mathrm{Tr}_{1\ldots N}\int_0^t dt'\left[\breve{V}(t),\left[\breve{V}(t'),\breve{\rho}(0)\right]\right]+\ldots \qquad (5.6.8)$$

Generally the derivative in Equation 5.6.8 is taken to be the "course-grain derivative" which means that the times involved are longer than the correlation times of the reservoir.

Finally for this topic, the interaction representation of the density operator at t=0 can be replaced by the Schrodinger representation since

$$\breve{\rho}(t)=\hat{u}^+(t)\,\hat{\rho}(t)\,\hat{u}(t)\qquad\text{with}\qquad\hat{u}=\exp\left(\frac{\hat{H}_o t}{i\hbar}\right)$$

Setting t=0 provides

$$\hat{u}(0)=1\qquad\text{and therefore}\qquad\breve{\rho}(0)=\hat{u}^+(0)\,\hat{\rho}(0)\,\hat{u}(0)=\hat{\rho}(0)$$

Equation 5.6.8 becomes

$$\dot{\breve{\rho}}_s=\frac{1}{i\hbar}\mathrm{Tr}_{1\ldots N}\left[\breve{V},\rho(0)\right]+\frac{1}{(i\hbar)^2}\mathrm{Tr}_{1\ldots N}\int_0^t dt'\left[\breve{V}(t),\left[\breve{V}(t'),\rho(0)\right]\right]+\ldots \qquad (5.6.9)$$

Topic 5.6.4: The Langevin Displacement Term

The first term in Equation 5.6.9 gives

$$\mathrm{Tr}_{1\ldots N}\left[\breve{V},\rho(0)\right]=\mathrm{Tr}_{1\ldots N}\sum_{ai}\left[\breve{S}_i^{(a)}\breve{R}_i^{(a)},\rho(0)\right]=\mathrm{Tr}_{1\ldots N}\sum_{ai}\left[\breve{S}_i^{(a)}\breve{R}_i^{(a)},\rho^{(s)}(0)\rho^{(1\ldots N)}(0)\right]$$

$$=\mathrm{Tr}_{1\ldots N}\sum_{ai}\left\{\rho^{(s)}(0)\breve{S}_i^{(a)}\left[\breve{R}_i^{(a)},\rho^{(1\ldots N)}(0)\right]+\left[\breve{S}_i^{(a)},\rho^{(s)}(0)\right]\breve{R}_i^{(a)}\rho^{(1\ldots N)}(0)\right\}$$

The trace refers to the reservoir degrees of freedom. The terms in the last equation are all similar.

$$\mathrm{Tr}_{1\ldots N}\breve{R}_i^{(a)}\rho^{(1\ldots N)}(0)=\mathrm{Tr}_1\rho^{(1)}(0)\ldots\mathrm{Tr}_{a-1}\rho^{(a-1)}(0)\quad\mathrm{Tr}_a\breve{R}_i^{(a)}\rho^{(a)}(0)\quad\mathrm{Tr}_{a+1}\rho^{(a+1)}(0)\ldots\mathrm{Tr}_N\rho^{(N)}(0) \qquad (5.6.10)$$

Factors without the reservoir variable $\breve{R}_i^{(a)}$ are evaluated similar to

$$\mathrm{Tr}_1\rho^{(1)}(0)=1$$

by definition of the density operator. The factors with the reservoir variable give

$$\mathrm{Tr}_a\breve{R}_i^{(a)}\rho^{(a)}(0)=\mathrm{Tr}_a\rho^{(a)}(0)\breve{R}_i^{(a)}=\left\langle\breve{R}_i^{(a)}\right\rangle_{\substack{res\\a}}=0$$

since the $\hat{R}$ are either creation or annihilation operators and $\langle n|\hat{R}|n\rangle\sim\langle n|n\pm1\rangle=0$. Also notice that the cyclic property of the trace was used to rearrange the order of "R" and "ρ" as necessary. Therefore, the Langevin displacement term is

$$Tr_{1\cdots N}\left[\tilde{V},\rho(0)\right]=0$$

where the last equality follows from the cyclic property of the trace operation.

Taking the trace over the reservoir states is equivalent to finding the ensemble average of $\hat{R}_i$. The ensemble average is equivalent to an average taken over a sufficiently long time interval (for an ergodic process). This is where the question of the interval of integration becomes important and why the time derivative in Equation 5.6.7 is usually called the "course grain derivative". If the values of the random variables fluctuate and are correlated over a small time interval, then the time average of the random variable might depend on the length of the time interval. The figure for example shows that the average of some variable R over the time interval [0,1] is nonzero while over [0,2] it is zero. Presumably, with long enough integration time, the average is zero. This time average (over sufficiently long times) matches the ensemble average since every duplicate system in the ensemble is in a different possible state. These considerations are important for the course-grain derivative in Equation 5.6.7. The time interval Δt is assumed to be longer than any correlation time of the reservoir. These comments apply to any of the integrals over time. In this way, the average values can be replaced with a definite value. The mutliple reservoir theory assumes that the average of the operators $\hat{R}$ is zero; however, the zero-value of $\langle\hat{R}\rangle_{res}$ is rigorously true when $\hat{R}_i$ are creation and annihilation operators. Quantum mechanical expectation values of creation or annihilation operators provide $\langle n|\hat{b}|n\rangle=0$.

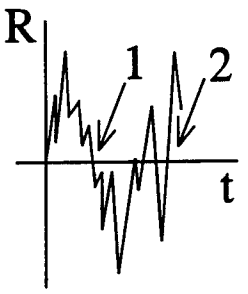


Figure 5.6.1: The average of R over the interval (0,2) is zero whereas it is nonzero on (0,1).

Topic 5.6.5: The Relaxation Term

The second term in Equation 5.6.9 gives the relaxation effects

$$\text{Term 2}=Tr_{1\cdots N}\int_0^t dt'\left[\tilde{V}(t),\left[\tilde{V}(t'),\rho(0)\right]\right]$$

Expanding the commutators provide

$$\text{Term 2}=Tr_{1\cdots N}\int_0^t dt'\left[\tilde{V}(t),\tilde{V}(t')\rho(0)-\rho(0)\tilde{V}(t')\right]$$

$$=Tr_{1\cdots N}\int_0^t dt'\left[\begin{array}{c}\tilde{V}(t)\tilde{V}(t')\rho(0)-\tilde{V}(t)\rho(0)\tilde{V}(t')-\tilde{V}(t')\rho(0)\tilde{V}(t)+\rho(0)\tilde{V}(t')\tilde{V}(t)\\ \leftarrow\text{Term 2.1}\rightarrow\ \ \leftarrow\text{Term 2.2}\rightarrow\ \ \leftarrow\text{Term 2.3}\rightarrow\ \ \leftarrow\text{Term 2.4}\rightarrow\end{array}\right] \tag{5.6.11}$$

Examine each sub-term in Equation 5.6.11. Substitute expressions for the interaction Hamiltonian in each one.

$$\text{Term 2.1} = \text{Tr}_{1\ldots N}\int_0^t dt'\,\breve{V}(t)\breve{V}(t')\rho(0) = \text{Tr}_{1\ldots N}\int_0^t dt'\sum_{ai}\breve{S}_i^{(a)}(t)\breve{R}_i^{(a)}(t)\sum_{bj}\breve{S}_j^{(b)}(t')\breve{R}_j^{(b)}(t')\rho(0)$$

$$= \text{Tr}_{1\ldots N}\int_0^t dt'\sum_{\substack{ij\\ab}}\breve{S}_i^{(a)}(t)\breve{R}_i^{(a)}(t)\breve{S}_j^{(b)}(t')\breve{R}_j^{(b)}(t')\rho^{(s)}(0)\,\rho^{(1\ldots N)}(0)$$

where the summation was moved to the left and the expression for the density operator was substituted. Moving the trace inside the summation and commuting system operators with reservoir operators provides

$$\text{Term 2.1} = \int_0^t dt'\sum_{\substack{ij\\ab}}\breve{S}_i^{(a)}(t)\breve{S}_j^{(b)}(t')\rho^{(s)}(0)\,\text{Tr}_{1\ldots N}\left[\breve{R}_i^{(a)}(t)\breve{R}_j^{(b)}(t')\rho^{(1\ldots N)}(0)\right]$$

Following the procedure outlined in the previous subtopic (see Equation 5.6.10), the trace factor can be rewritten to provide

$$\text{Term 2.1} = \int_0^t dt'\sum_{\substack{ij\\ab}}\breve{S}_i^{(a)}(t)\breve{S}_j^{(b)}(t')\rho^{(s)}(0)\left\langle\breve{R}_i^{(a)}(t)\breve{R}_j^{(b)}(t')\right\rangle\delta_{ab}$$

$$= \int_0^t dt'\sum_{ija}\breve{S}_i^{(a)}(t)\breve{S}_j^{(a)}(t')\rho^{(s)}(0)\left\langle\breve{R}_i^{(a)}(t)\breve{R}_j^{(a)}(t')\right\rangle$$

since for $a \neq b$ the fluctuations average to zero

$$\text{Tr}_{1\ldots N}\left[\breve{R}_i^{(a)}(t)\breve{R}_j^{(b)}(t')\rho^{(1\ldots N)}(0)\right] = \left\langle\breve{R}_i^{(a)}(t)\right\rangle\left\langle\breve{R}_j^{(b)}(t')\right\rangle = 0$$

Similar reasoning applies to the other parts of "term 2".

$$\text{Term 2.2} = \int_0^t dt'\sum_{ija}\breve{S}_i^{(a)}(t)\rho^{(s)}(0)\,\breve{S}_j^{(a)}(t')\left\langle\breve{R}_j^{(a)}(t')\breve{R}_i^{(a)}(t)\right\rangle$$

$$\text{Term 2.3} = \int_0^t dt'\sum_{ija}\breve{S}_i^{(a)}(t')\rho^{(s)}(0)\,\breve{S}_j^{(a)}(t)\left\langle\breve{R}_j^{(a)}(t)\breve{R}_i^{(a)}(t')\right\rangle$$

$$\text{Term 2.4} = \int_0^t dt'\sum_{ija}\rho^{(s)}(0)\breve{S}_i^{(a)}(t')\breve{S}_j^{(a)}(t)\left\langle\breve{R}_i^{(a)}(t')\breve{R}_j^{(a)}(t)\right\rangle$$

Where the cyclic property of the trace is used to obtain the results for Terms 2.2 and 2.3. For terms 2.3 and 2.4, interchange the dummy indices "i" and "j". For all terms, the fluctuations are stationary so that, for example, we can write

$$\left\langle\breve{R}_i^{(a)}(t')\breve{R}_j^{(a)}(t)\right\rangle = \left\langle\hat{R}_i^{(a)}\breve{R}_j^{(a)}(t-t')\right\rangle$$

We therefore find

$$\text{Term 2.1} = \int_0^t dt'\sum_{ija}\breve{S}_i^{(a)}(t)\breve{S}_j^{(a)}(t')\rho^{(s)}(0)\left\langle\breve{R}_i^{(a)}(t-t')\hat{R}_j^{(a)}\right\rangle$$

$$\text{Term 2.2} = \int_0^t dt'\sum_{ija}\breve{S}_i^{(a)}(t)\rho^{(s)}(0)\,\breve{S}_j^{(a)}(t')\left\langle\hat{R}_j^{(a)}\breve{R}_i^{(a)}(t-t')\right\rangle$$

$$\text{Term 2.3} = \int_0^t dt'\sum_{ija}\breve{S}_j^{(a)}(t')\rho^{(s)}(0)\,\breve{S}_i^{(a)}(t)\left\langle\breve{R}_i^{(a)}(t-t')\hat{R}_j^{(a)}\right\rangle$$

$$\text{Term 2.4} = \int_0^t dt' \sum_{ija} \rho^{(s)}(0)\breve{S}_j^{(a)}(t')\breve{S}_i^{(a)}(t)\left\langle \hat{R}_j^{(a)}\breve{R}_i^{(a)}(t-t')\right\rangle$$

Changing the variable of integration for all terms from t' to $\tau = t - t'$ provides

$$\text{Term 2.1} = \int_0^t dt \sum_{ija} \breve{S}_i^{(a)}(t)\breve{S}_j^{(a)}(t-\tau)\rho^{(s)}(0)\left\langle \breve{R}_i^{(a)}(\tau)\hat{R}_j^{(a)}\right\rangle$$

$$\text{Term 2.2} = \int_0^t d\tau \sum_{ija} \breve{S}_{ij}^{(a)}(t)\rho^{(s)}(0)\breve{S}_j^{(a)}(t-\tau)\left\langle \hat{R}_j^{(a)}\breve{R}_i^{(a)}(\tau)\right\rangle$$

$$\text{Term 2.3} = \int_0^t d\tau \sum_{ija} \breve{S}_j^{(a)}(t-\tau)\rho^{(s)}(0)\breve{S}_i^{(a)}(t)\left\langle \breve{R}_i^{(a)}(\tau)\hat{R}_j^{(a)}\right\rangle$$

$$\text{Term 2.4} = \int_0^t d\tau \sum_{ija} \rho^{(s)}(0)\breve{S}_j^{(a)}(t-\tau)\breve{S}_i^{(a)}(t)\left\langle \hat{R}_j^{(a)}\breve{R}_i^{(a)}(\tau)\right\rangle$$

The sub-terms can all be recombined into

$$\dot{\breve{\rho}}^{(s)}(t) = 0 + \text{Term 2} = \text{Tr}_{1\cdots N}\int_0^t dt'\left[\breve{V}(t),\left[\breve{V}(t'),\rho(0)\right]\right]$$

The density operator evolves according to

$$\dot{\breve{\rho}}^{(s)}(t) = -\frac{1}{\hbar^2}\sum_{ija}\int_0^\infty d\tau\,\{\left[\breve{S}_i^{(a)}(t),\breve{S}_j^{(a)}(t-\tau)\breve{\rho}^{(s)}(t)\right]\left\langle \breve{R}_i^{(a)}(\tau)\hat{R}_j^{(a)}\right\rangle + \\ -\left[\breve{S}_i^{(a)}(t),\breve{\rho}^{(s)}(t)\breve{S}_j^{(a)}(t-\tau)\right]\left\langle \hat{R}_j^{(a)}\breve{R}_i^{(a)}(\tau)\right\rangle\} \tag{5.6.12}$$

The limit "t" on the integral is replaced by ∞ since the correlation functions are non-zero only for very small times. Equation 5.6.12 also contains the Markov approximation

$$\hat{\rho}^{(s)}(0) = \breve{\rho}^{(s)}(0) \to \breve{\rho}^{(s)}(t)$$

Topic 5.6.6: The Pauli Master Equation

The development in this topic follows Weissbluth's approach for a single reservoir; however, here multiple reservoirs are included. Let $\{|k\rangle\}$ be eigenvectors of $\hat{H}_s$, the system Hamiltonian.

$$\dot{\breve{\rho}}^{(s)}(t) = -\frac{1}{\hbar^2}\sum_{ija}\int_0^\infty d\tau\,\{\underbrace{\left[\breve{S}_i^{(a)}(t),\breve{S}_j^{(a)}(t-\tau)\breve{\rho}^{(s)}(t)\right]\left\langle \breve{R}_i^{(a)}(\tau)\hat{R}_j^{(a)}\right\rangle}_{\leftarrow \text{Term 3.1} \rightarrow} \underbrace{-\left[\breve{S}_i^{(a)}(t),\breve{\rho}^{(s)}(t)\breve{S}_j^{(a)}(t-\tau)\right]\left\langle \hat{R}_j^{(a)}\breve{R}_i^{(a)}(\tau)\right\rangle}_{\leftarrow \text{Term 3.2} \rightarrow}\} \tag{5.6.13}$$

After considerable algebraic manipulation, Term 3.1 in $\dot{\breve{\rho}}^{(s)}(t)$ can be rewritten using the closure relation for the system eigenstates

$$\langle k|\left[\breve{S}_i^{(a)}(t),\breve{S}_j^{(a)}(t-\tau)\breve{\rho}^{(s)}(t)\right]|l\rangle = \sum_{nm}\langle m|\breve{\rho}^{(s)}(t)|n\rangle\left\{\delta_{ln}\sum_r\langle k|S_i^{(a)}|r\rangle\langle r|S_j^{(a)}|m\rangle e^{-i\omega_{rm}t} + \right. \\ \left. -\langle n|S_i^{(a)}|l\rangle\langle k|S_j^{(a)}|m\rangle e^{-i\omega_{km}t}\right\}e^{i(\omega_{km}+\omega_{nl})t}$$

Term 3.2 becomes

$$\langle k|\left[\breve{S}_i^{(a)}(t),\breve{\rho}^{(s)}(t)\breve{S}_j^{(a)}(t-\tau)\right]|l\rangle=\sum_{nm}\langle m|\breve{\rho}^{(s)}(t)|n\rangle\left\{\langle n|S_j^{(a)}|l\rangle\langle k|S_i^{(a)}|m\rangle e^{-i\omega_{nl}t}+\right.$$
$$\left.-\delta_{km}\sum_r\langle n|S_j^{(a)}|r\rangle\langle r|S_i^{(a)}|l\rangle e^{-i\omega_{nr}t}\right\}e^{i(\omega_{km}+\omega_{nl})t}$$

where "l" is lower case "L". The difference in angular frequency is

$$\omega_{nl}=\frac{1}{\hbar}(E_n-E_l)$$

where E_n (etc.) refers to the energy of n^{th} level of the system. Notice that ω_{nl} can be negative.

Define matrix elements

$$\Gamma_{nlkm}^{(a)+}=\frac{1}{\hbar^2}\sum_{ij}\langle n|S_i^{(a)}|l\rangle\langle k|S_j^{(a)}|m\rangle\int_0^\infty d\tau e^{-i\omega_{km}\tau}\left\langle\breve{R}_i^{(a)}(\tau)R_j^{(a)}\right\rangle$$

$$\Gamma_{nlkm}^{(a)-}=\frac{1}{\hbar^2}\sum_{ij}\langle n|S_j^{(a)}|l\rangle\langle k|S_i^{(a)}|m\rangle\int_0^\infty d\tau e^{-i\omega_{nl}\tau}\left\langle R_j^{(a)}\breve{R}_i^{(a)}(\tau)\right\rangle$$

and also the relaxation matrix elements

$$R_{klmn}^{(a)}=-\delta_{ln}\sum_r\Gamma_{krrm}^{(a)+}+\Gamma_{nllm}^{(a)+}+\Gamma_{nlkm}^{(a)-}-\delta_{km}\sum_r\Gamma_{nrrl}^{(a)-}$$

Substitute all five of the previous expressions into Equation 5.6.13 to find

$$\langle k|\dot{\breve{\rho}}_s(t)|l\rangle=\sum_{nma}\langle m|\breve{\rho}_s(t)|n\rangle R_{klmn}^{(a)}e^{i(\omega_{km}+\omega_{nl})t} \qquad (5.6.14)$$

Applying the rotating wave approximation (RWA) requires

$$\omega_{km}+\omega_{nl}=0$$

There are three cases implied by the RWA

(k=m, l=n, $k\neq l$) (k=l, m=n, $k\neq n$) (k=l=n=m)

From these, two separate equations can be found

$$\langle k|\dot{\breve{\rho}}_s(t)|l\rangle=\sum_a\langle k|\breve{\rho}_s(t)|l\rangle R_{klkl}^{(a)}$$

$$\langle k|\dot{\breve{\rho}}|k\rangle=\sum_a\sum_{\substack{n\\n\neq k}}\langle n|\breve{\rho}_s(t)|n\rangle R_{kknn}^{(a)}$$

Define

$$R_{ijkl}=\sum_a R_{ijkl}^{(a)}$$

and combine the previous two equations to obtain

$$\langle k|\dot{\breve{\rho}}_s(t)|l\rangle=\langle k|\breve{\rho}_s(t)|l\rangle R_{klkl}+\delta_{kl}\sum_{\substack{n\\n\neq k}}\langle n|\breve{\rho}_s(t)|n\rangle R_{kknn} \qquad (5.6.15)$$

Notice that the effect of all the reservoirs combines into a single constant. This equation clearly show that the reservoir causes a relaxation of the density operator (the system changes state due to the reservoir interaction).

Equation 5.6.15 shows how the density operator changes due to the interaction between the reservoir and system. However, the equation does not explicitly show the change in the density operator due to interaction within the system (such as interaction

between the system atom and light). The equation that describes the total motion of the density operator can be demonstrated by using the equation of motion for the system density operator in the Schrodinger representation. Equation 5.3.12 in Topic 5.3.4 is repeated here as

$$\frac{\partial \hat{\rho}_s(t)}{\partial t} = \hat{u}_s \dot{\tilde{\rho}}_s(t) \hat{u}_s^+ + \frac{1}{i\hbar}\left[\hat{H}_s, \hat{\rho}_s(t)\right] \tag{5.6.16}$$

To combine Equations 5.6.15 and 5.6.16, requires that we take the "kl" matrix element of Equation 5.6.16.

$$\begin{aligned} \left.\frac{\partial \hat{\rho}_s(t)}{\partial t}\right]_{kl} &= \left\langle k\left|\left[\hat{u}_s \dot{\tilde{\rho}}_s(t) \hat{u}_s^+\right]\right|l\right\rangle + \frac{1}{i\hbar}\left[\hat{H}_s, \hat{\rho}_s(t)\right]_{kl} \\ &= \left\langle k\left|\left[e^{\frac{\hat{H}_s t}{i\hbar}} \dot{\tilde{\rho}}_s(t) e^{-\frac{\hat{H}_s t}{i\hbar}}\right]\right|l\right\rangle + \frac{1}{i\hbar}\left[\hat{H}_s, \hat{\rho}_s(t)\right]_{kl} \end{aligned} \tag{5.6.17}$$

Operating with the unitary operators inside the first expectation value on the right-hand side provides

$$\left\langle k\left|\left[e^{\frac{\hat{H}_s t}{i\hbar}} \dot{\tilde{\rho}}_s(t) e^{-\frac{\hat{H}_s t}{i\hbar}}\right]\right|l\right\rangle = e^{i\omega_{lk}t}\left\langle k\left|\dot{\tilde{\rho}}_s(t)\right|l\right\rangle$$

so that Equation 5.6.161 becomes

$$\left.\frac{\partial \hat{\rho}_s(t)}{\partial t}\right]_{kl} = e^{i\omega_{lk}t}\left\langle k\left|\dot{\tilde{\rho}}_s(t)\right|l\right\rangle + \frac{1}{i\hbar}\left[\hat{H}_s, \hat{\rho}_s(t)\right]_{kl} \tag{5.6.18}$$

Combining Equations 5.6.15 and 5.6.18 provides

$$\left.\frac{\partial \hat{\rho}_s(t)}{\partial t}\right]_{kl} = \frac{1}{i\hbar}\left[\hat{H}_s, \hat{\rho}_s(t)\right]_{kl} + e^{i\omega_{lk}t}\left\langle k\left|\tilde{\rho}_s(t)\right|l\right\rangle R_{klkl} + e^{i\omega_{lk}t}\delta_{kl}\sum_{\substack{n \\ n\neq k}}\left\langle n\left|\tilde{\rho}_s(t)\right|n\right\rangle R_{klnn}$$

Writing the interaction density operator in terms of the Schrodinger operator, yields

$$\left.\frac{\partial \hat{\rho}_s(t)}{\partial t}\right]_{kl} = \frac{1}{i\hbar}\left[\hat{H}_s, \hat{\rho}_s(t)\right]_{kl} + e^{i\omega_{lk}t}\left\langle k\left|\hat{u}_s^+ \hat{\rho}_s(t) \hat{u}_s\right|l\right\rangle R_{klkl} + e^{i\omega_{lk}t}\delta_{kl}\sum_{\substack{n \\ n\neq k}}\left\langle n\left|\hat{u}_s^+ \hat{\rho}_s(t) \hat{u}_s\right|n\right\rangle R_{klnn}$$

The equation can be rewritten by substituting the exponential form of the unitary operators

$$\hat{u}_s = \exp\left(\frac{\hat{H}_s t}{i\hbar}\right)$$

which operate on the basis vectors. The operation yields a factor of $e^{-i\omega_{lk}t}$ to yield

$$\frac{\partial \hat{\rho}_{kl}}{\partial t} = \frac{1}{i\hbar}\left[\hat{H}_s, \hat{\rho}_s(t)\right]_{kl} + e^{i\omega_{lk}t}e^{-i\omega_{lk}t}\left\langle k\left|\hat{\rho}_s(t)\right|l\right\rangle R_{klkl} + e^{i\omega_{lk}t}\delta_{kl}\sum_{\substack{n \\ n\neq k}} e^{i\omega_n t - i\omega_n t}\left\langle n\left|\hat{\rho}_s(t)\right|n\right\rangle R_{klnn}$$

Canceling exponential terms and noting that the factor δ_{kl} results in $\omega_{lk} = \omega_l - \omega_k = 0$, we end up with

$$\frac{\partial \hat{\rho}_{kl}^{(s)}}{\partial t} = \frac{1}{i\hbar}\left[\hat{H}_s, \hat{\rho}^{(s)}(t)\right]_{kl} + \left\langle k\left|\hat{\rho}^{(s)}(t)\right|l\right\rangle R_{klkl} + \delta_{kl}\sum_{\substack{n \\ n\neq k}}\left\langle n\left|\hat{\rho}^{(s)}(t)\right|n\right\rangle R_{klnn} \tag{5.6.19}$$

Equation 5.6.19 is the Liouville equation for the density matrix as discussed in the first several sections of this chapter. The commutator provides the dynamics internal to the system itself. As a very important point, the exact composition of the "system" is unspecified; it can be composed of an atom and an electromagnetic wave interacting with the atom. Using notation previously employed, the last two terms in Equation 5.6.19, can be written as

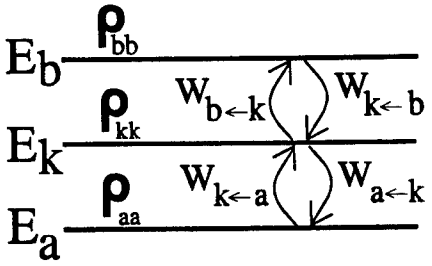


Figure 5.6.2: The rates of transition "W" between energy levels "E". The diagonal density matrix elements give the population of each level.

$$\left.\frac{\partial \hat{\rho}^{(s)}_{kl}}{\partial t}\right|_{other} = \langle k|\hat{\rho}^{(s)}(t)|l\rangle R_{klkl} + \delta_{kl}\sum_{\substack{n \\ n\neq k}}\langle n|\hat{\rho}^{(s)}(t)|n\rangle R_{kknn}$$

The rate of change of the diagonal elements are

$$\left.\frac{\partial \hat{\rho}^{(s)}_{kk}}{\partial t}\right|_{other} = \langle k|\hat{\rho}^{(s)}(t)|k\rangle R_{kkkk} + \sum_{\substack{n \\ n\neq k}}\langle n|\hat{\rho}^{(s)}(t)|n\rangle R_{kknn}$$ The first term is the transition out of level "k" and the second term is the transition rate into level "k" from other levels "n".

Chapter 6: The Wavelength-Dependent Transfer Function for Optically-Controlled Semiconductor Lasers

The optically controlled laser, consisting of a main semiconductor laser with an associated control beam, is important for integrated high-speed optical signal processing and for optical communications. The optically-controlled laser rate equations, the gain, and the wavelength dependent transfer function are derived from the Liouville Equation for the density matrix describing a system with multiple perturbing optical fields. The results explain the primary effects of coherent and non-coherent control signals upon the output of the main laser (illustrated with experimental data). New predictions are obtained for optical gain, wavelength conversion and FM to AM demodulation. This chapter is a reprint of the Parker et. al. publication in J. Appl. Phys. 83, 5056 (1998).

Section 6.1 Introduction

Lasers with optical gain control are important due to the possibility of integrated all-optical high-speed communications and signal processing. Recent publications discuss optically-controlled lasers that are capable of signal routing and logic functions. These reported devices have a variety of configurations but generally, they consist of a semiconductor main laser with a control beam entering the side of the cavity at right angles (refer to Figure 6.1). The main laser can be an in-plane laser (IPL) or a vertical cavity laser (VCSEL). For the reported GaAs-based devices, the source of the control beam is another semiconductor laser monolithically integrated with the main laser. However, the control beam can also originate in a monolithically integrated laser amplifier or a fiber/waveguide that is butt-coupled to the side of the main laser. In operation, an increase of the control-beam power P_c (with angular frequency ω_c) causes the output-beam power P (with angular frequency ω) to decrease. Similar optical modulation effects are seen for fiber lasers using DBR mirrors and two longitudinal modes.

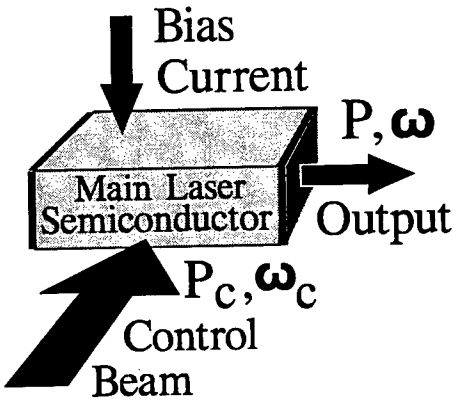


Figure 6.1: Block diagram of a Laser with Optical Gain Control. The semiconductor forms the cavity of "main laser" which can be either an edge-emitting laser or a VCSEL.

Although these publications show the operation of the optically-controlled lasers, many features remain to be explained. Both bistable and linear switching characteristics are observed. Experimentally, three distinct regions of non-bistable operation are reported for devices with integrated control lasers (see the top curve in Figure 6.2). The spontaneous emission region (SER) corresponds to low values of *bias current* I_c applied to the integrated *control laser*. The slope $\frac{\partial P}{\partial I_c}$ is relatively small but it can be positive, negative or zero. The SER describes the transfer function when the *control* laser is

operated below threshold; apparently, spontaneous emission is ineffective for quenching or pumping the main laser. In some cases, the *main-laser threshold current* I_{thr} is observed to decrease as I_c increases. The linear region (LR), as shown for mid-range values of bias current, has a slope $\frac{\partial P}{\partial I_c}$ (and also $\frac{\partial P}{\partial P_c}$) that is relatively constant and negative; this region marks one of the most distinctive features of the optically-controlled laser (see also curves Q and H in Figure 6.2). For the LR, the main-laser threshold current I_{thr} is observed to increase linearly as P_c increases. Of the three regions, the LR has the largest magnitude for the optical gain as defined by $G_{opt} \equiv \left|\frac{dP}{dP_c}\right|$. Interestingly, the values of the optical gain have never been observed to be larger than one. For the saturation region (SR), increasing the control laser bias current results in little change of the main-laser output power ($\frac{\partial P}{\partial I_c}$ and $\frac{\partial P}{\partial P_c}$ are nearly zero). Often the magnitude of the slope of the SR rapidly decreases and the curve approaches a clear nonzero asymptote (within a span of 10 to 20 mA for I_c).

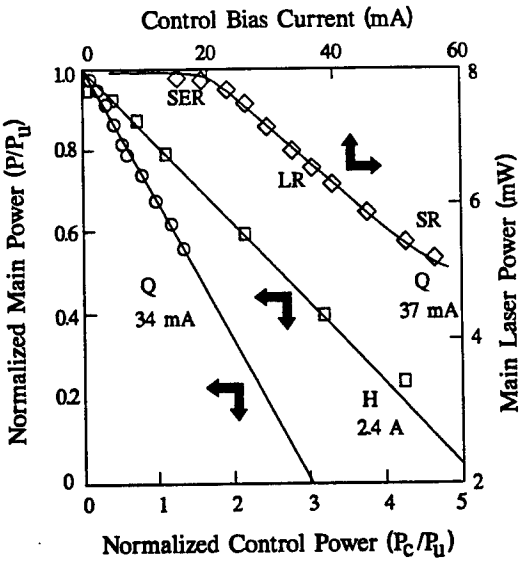


Figure 6.2: Experimental transfer characteristics for (Q) an edge-emitting multiple quantum well laser with ridge waveguiding and (H) a homojunction laser (see Ref. 7,8). The bias current to the quanum well laser is 34 mA while that to the homojunction laser is 2.4 Amps. The power P from the main laser and the power P_c of the control beam are normalized by the "unquenched" power P_u for the main-laser. The top curve (a P-I_c plot) refers to the Q laser (37 mA bias current) and shows three regions of non-bistable operation.

Of primary importance for understanding these observations and applying the optically-controlled laser is the transfer function that describes the main laser output power P as a function of the main laser bias current I and angular frequency ω, the control source power P_c, the control beam wavelength λ_c (angular frequency ω_c), and the device geometry. In section 6.2, we derive the wavelength-dependent quenched-laser rate equations and material gain. The transfer function is derived in Section 6.3. The theory is applied in Section 6.4 to the observed operating characteristics. Besides signal routing and optical logic functions, the optically controlled lasers are desirable for low-noise optical amplification, wavelength conversion and FM demodulation as will be discussed in Section 6.5.

6.2 The Rate Equations and Gain for Optically Controlled Lasers

We derive the rate and gain equations (for lasers with optical gain control) with careful attention to the initial detail since the theory has never been applied to optically controlled lasers and since it provides the framework for understanding and applying the

devices. In addition, we highlight the assumptions (so that the work can be generalized in the future) and indicate where the derivation differs from standard treatments (such as those in the references).

For the optically controlled lasers, there are at least two optical fields $\vec{F}_j = \frac{\vec{\varepsilon}_j}{2}\left[e^{i\omega_j t} + e^{-i\omega_j t}\right]$ interacting with the gain medium of the laser. The first (j=1) and second (j=2), respectively, refer to the main laser field and the control field (which fully overlaps the main cavity). These fields are assumed to interact with electrons in conduction band (cb) and valence band (vb) Bloch eigenstates $\left|E,\vec{k}\right\rangle$ of the unperturbed Hamiltonian H_o (E is the energy and $\vec{k}$ is the Bloch wavevector restricted to the first Brillouin zone). The medium is electrically (or optically) pumped to maintain the excess carrier population that interacts with the applied fields.

The Liouville equation for the density matrix[9-11] is

$$\frac{\partial \rho_{MN}}{\partial t} = \frac{-i}{\hbar}[H,\rho]_{MN} - \frac{\rho_{MN} - \bar{\rho}_{MN}}{\tau_{MN}} \qquad (6.1)$$

where τ_{MN} is the relaxation (diagonal terms) and dephasing time (off-diagonal terms), $\rho_{M,N} \equiv \left\langle E_m \vec{k}_a \left| \rho \right| E_n \vec{k}_b \right\rangle$ (where m,n are in {1,2}, and a,b label the allowed $\vec{k}$), $\bar{\rho}_{MN}$ is defined below, $H=H_o+V$ is the Hamiltonian, $V = \sum_j \mu^j F_j$ is the time dependent perturbation for the transistions,[12] μ^j is the transition dipole operator acting along the polarization direction of $\vec{F}_j$. We assume for semiconductor lasers that the dipole operator μ^j and the perturbation V are diagonal[12] in the Bloch wave-vector $\vec{k}$. The matrices of μ^j and V therefore reduce to $\bigoplus_{\vec{k}} \mu^{j(\vec{k})}_{mn}$ and $\bigoplus_{\vec{k}} V^{(\vec{k})}_{mn}$ which consist of 2x2 matrices (m,n in {1,2}) arranged on the diagonal of a large matrix that is indexed[9] by the wave vectors $\vec{k}$. On the time scale of interest, we assume equilibrium between states within the same band. The quasi-equilibrium (Q-E) density matrix $\bar{\rho}_{MN}$ is taken to be diagonal in both E and $\vec{k}$ with each entry equal to the probability of an electron occupying a state in the cb or vb in the absence of the perturbation according to $f(E) = \left\{1 + \exp\left[(E - F_{c,v})/kT\right]\right\}^{-1}$. The diagonal Q-E density matrix reduces[11] to $\left[\bar{\rho}\right] = \bigoplus_{\vec{k}} \begin{bmatrix} f(E_2^{(\vec{k})}) & 0 \\ 0 & f(E_1^{(\vec{k})}) \end{bmatrix}$ where the energy of a cb electron $E_2^{(\vec{k})}$ and vb electron $E_1^{(\vec{k})}$ depend on the band structure.[12] The density matrix therefore reduces to $\rho_{MN} = \bigoplus_{\vec{k}} \rho^{(\vec{k})}_{mn}$ (for m,n in {1,2}).[13] As a key result, Equation 1 applies to each 2x2 submatrix $\rho^{(\vec{k})}_{mn}$ and Equation 1 can be recast into a set of independent equations, one set for each different $\vec{k}$.

The Liouville equation becomes a set of equations similar to those for a collection of two-level atoms except that each quantity has a field index j and Bloch wavevector index $\vec{k}$.

$$\frac{\partial \rho_{12}^{j(\vec{k})}}{\partial t} = \frac{i}{\hbar}\left[\hbar\omega_o \rho_{12}^{j(\vec{k})} - V_{12}^{(\vec{k})}\eta^{(\vec{k})}\right] - \frac{\rho_{12}^{j(\vec{k})}}{\tau_2} \tag{6.2}$$

$$\frac{\partial \eta^{(\vec{k})}}{\partial t} = \frac{-2i}{\hbar} V_{12}^{(\vec{k})}\left[\rho_{12}^{j(\vec{k})} - \rho_{12}^{j(\vec{k})^+}\right] - \frac{\eta^{(\vec{k})} - \bar{\eta}^{(\vec{k})}}{\tau} \tag{6.3}$$

where $\rho_{12}^{(\vec{k})} = \sum_j \rho_{12}^{j(\vec{k})}$, and $\bar{\eta}^{(\vec{k})} = \bar{\rho}_{22}^{(\vec{k})} - \bar{\rho}_{11}^{(\vec{k})} = f_c(E_2^{(\vec{k})}) - f_v(E_1^{(\vec{k})})$. As an important point for multiple fields, the terms related to the induced polarization ($\rho_{12}^{j(\vec{k})} = \rho_{21}^{j(\vec{k})^+}$) are indexed by the field j (where the superscript "+" denotes conjugation) but the difference between the occupation probabilities $\eta^{(\vec{k})} = \rho_{22}^{(\vec{k})} - \rho_{11}^{(\vec{k})}$ is not indexed. The sum over j (i.e. ω_j) for $\rho_{mn}^{(\vec{k})}$ is a Fourier decomposition whereas $\eta^{(\vec{k})}$ can be shown to have a "DC" Fourier component. The steady-state difference in occupation probability $\eta^{(\vec{k})}$ in Equation 3 can be independent of time (i.e., it does not need a term proportional to $e^{\pm i\omega_j t}$) since $\rho_{12}^{j(\vec{k})}$ has a time dependent component to cancel out the appropriate time dependent part of V_{12} (and assuming the rotating-wave approximation). Equation 2 is satisfied under these conditions since each term in the equation will have components proportional to $e^{\pm i\omega_j t}$.

The equivalence between the equations for a collection of two level atoms and for the quenched laser is exploited to avoid some algebra (see Reference 10). In the steady-state, expressions that satisfy the modified optical Bloch Equations (Equations 6.2,6.3) are given by

$$\eta^{(\vec{k})} = \frac{\bar{\eta}^{(\vec{k})}}{1 + \sum_j 2\Omega_j^{(\vec{k})2}\tau L_j^{(\vec{k})}} \tag{6.4}$$

$$\rho_{12}^{j(\vec{k})} = -\frac{\eta^{(\vec{k})}\Omega_j^{(\vec{k})}}{D_j^{(\vec{k})}}\left[i\tau_2 + (\omega_j - \omega_o^{(\vec{k})})(\tau_2)^2\right]e^{i\omega_j t} \tag{6.5}$$

where $\Omega_j^{(\vec{k})} = \mu_j^{(\vec{k})}\varepsilon_j/(2\hbar)$, $\mu_j^{(\vec{k})} \equiv \mu_{12}^{j(\vec{k})} = \mu_{21}^{j(\vec{k})}$, $D_j^{(\vec{k})} = 1 + (\omega_j - \omega_o^{(\vec{k})})(\tau_2)^2$, $\hbar\omega_o^{(k)}$ is the transition energy at Bloch wave vector $\vec{k}$, and $L^{(\vec{k})}(\omega) = 2\tau_2 \Big/ \left[1 + (\omega - \omega_o^{(\vec{k})})^2(\tau_2)^2\right]$ is the lineshape function. We assume that the population relaxation time $\tau=\tau_{11}=\tau_{22}$ and the dephasing time $\tau_2=\tau_{12}=\tau_{21}$ are independent of $\vec{k}$.

The gain for each field is calculated as usual from the induced polarization P_j (for field j) except that each applied field j requires a summation over the Bloch wavevectors.

$$P_j = \text{trace}\left(\rho^j \mu^j\right) = \text{trace}\left(\bigoplus_{\vec{k}} \rho^{j(\vec{k})}\mu^{j(\vec{k})}\right) = \sum_{\vec{k}} \text{trace}\left(\rho^{j(\vec{k})}\mu^{j(\vec{k})}\right) \tag{6.6}$$

For a semiconductor, the summation is replaced by both an integral $\int_{E_g}^{\infty} dE_T$ and the reduced density of states $g_r(E_T)$, where E_g is the gap energy and E_T is the transition energy.[12] The material gain for multiple optical fields obtains from Equation 2 and 6:

$$G(\omega_j) = \frac{\omega_j \mu_j^2}{2cn\hbar\varepsilon_o} \int_{E_g}^{\infty} dE_T \eta(E_T) g_r(E_T) L(\omega_j - E_T/\hbar) \tag{6.7}$$

where E_T is substituted[12] for E_1 and E_2 by using the k-selection rule. The difference between the occupation probabilities is given by $\eta(E_T) = \rho_{22}(E_T) - \rho_{11}(E_T)$ and the difference between the quasi-equilibrium occupation-probabilities (i.e., the difference in occupation probability without any optical field) is $\bar{\eta}(E_T) = \bar{\rho}_{22}(E_T) - \bar{\rho}_{11}(E_T)$. The other factors are $L(\omega - E_T/\hbar) = 2\tau_2 \Big/ \left[1 + (\omega - E_T/\hbar)^2 (\tau_2)^2\right]$, $\eta(E_T) = \bar{\eta}(E_T) \Big/ \left[1 + \sum_i 2\Omega_i^2 \tau L(\omega_i - E_T/\hbar)\right]$, and $\Omega_i^2 = \left(\frac{\mu\varepsilon_i}{2\hbar}\right)^2 = \frac{\mu^2 I_i}{2\hbar^2 \varepsilon_o c^2 n}$ where the calculated intensity $I_i = P_i/A_i$ depends on the approximate cross-sectional area A_i of beam i (A_1=A and A_2=A_c are the cross-sectional areas of the main and control beams, respectively).

There are two control mechanisms for lasers with optical gain control. One mechanism, *gain saturation*, is related to the homogenous broadening evident in Equation 7. The gain $G(\omega_j)$ for beam j depends on the intensity of beam i through Ω_i in $\eta(E_T)$. A change in the control-beam intensity yields a change in the gain for the output beam and hence a change in the output power. If desired, the gain can be expanded in a Taylor series to provide the wavelength dependence of the self- and cross-saturation coefficients used to describe bistable operation.

The second control mechanism, *gain competition*, is described by a transfer function that obtains from the quenched-laser rate equations even when the saturation term in $\eta(E_T)$ is assumed to be negligible (i.e., Ω_i^2=0). The Liouville and gain equations produce the steady-state quenched-laser rate equations (as can be verified from the previous relations)

$$0 = -vG(\omega)P - vG(\omega_c)\eta_{cm} A P_c + \frac{v\alpha_m}{2e\Gamma}\frac{hc}{\lambda}\left[I - I_{thr}\right] \tag{6.8}$$

$$0 = vG(\omega)\Gamma P - \frac{P}{\tau_p} + \beta R_{sp} \tag{6.9}$$

where βR_{sp} is the rate of spontaneous emission coupled into the main-laser waveguide, v=c/n is the speed of light in material with refractive index n, and Γ is the confinement factor. Geometry factors are included in the coupling coefficient η_{cm} between the control and main laser and also in $A = \frac{\omega}{\omega_c}\frac{A_m}{A_c}R$ where $R \cong \alpha_m L_m/2 \cong (1\text{-}R_m)/2$ for a non-integrated control source and $R \cong \frac{\alpha_m}{\alpha_c} \cong \frac{1-R_m}{1-R_c}\frac{L_c}{L_m}$ for two monolithically integrated lasers with cross-coupled cavities (subscripts c,m refer to control and main lasers,

α_m=ln(R)/L is the mirror loss, L and R are the length and mirror reflectivity of the laser) and $\tau_p^{-1}=v(\alpha_m+\alpha_{int})$ where α_{int} is the usual internal loss (cm^{-1}). The main-laser bias current is $I=\frac{eV}{2\tau}\int_{E_g}^{\infty} dE\bar{\eta}(E)g_r(E)$ (where V is the volume of the active region). A similar definition holds for I_{thr} by replacing $\bar{\eta}$ with η. The factor of 2 occurs since η and $\bar{\eta}$ are the difference in electron density between the cb and vb (i.e., double the density of cb electrons by charge neutrality). Of course, the cavity-lifetime has the form $\tau_p=\tau_p(\omega)$ for frequency selective DBR mirrors.

6.3 The Steady-State Transfer Function

As the main result, Equation 6.9 approximately fixes the main-laser gain at $vG(\omega)\Gamma=1/\tau_p$ for steady-state operation above lasing threshold and Equation 8 then yields the transfer characteristics

$$P=\frac{\hbar\omega}{2e}\frac{\alpha_m}{\alpha_m+\alpha_{int}}[I-I_{thr}]-\eta_{cm}A\frac{G(\omega_c)}{G(\omega)}P_c \tag{6.10}$$

The condition on the main-laser gain ($vG(\omega)\Gamma=1/\tau_p$) determines the peak gain for $G(\omega)$ in Equation 6.7 which determines the quasi-Fermi levels $F_{c,v}$. Essentially, the condition of quasi-equilibrium requires the entire gain curve to slightly shift (up or down) by an amount that depends the shape of the bands. The bias current supplied to the main-laser replenishes the inverted carrier population and maintains the approximate peak gain of $G\cong(\tau_p v\Gamma)^{-1}$. With some approximation, Equation 6.10 can be modified for optically quenched lasers with partial overlap between the laser cavity and the control beam by replacing P_c with εP_c where ε is the ratio of the cross-sectional area of the control beam to the cross-sectional area of the laser cavity.[3]

The steady-state transfer characteristics for the optically-controlled lasers obtain from Equation 10. As an example, Figure 3 shows the ratio of the output- to the quiescent-power as a function of the photon-energy of the control beam. These transfer curves are obtained by rewriting the last equation as

$$\frac{P}{P_u}=1-\eta_{cm}A\left[\frac{G(\omega_c)}{G(\omega)}\right]\frac{P_c}{P_u} \tag{6.11}$$

where the unquenched power P_u is obtained from Equation 10 with P_c=0 (control beam "off"). The horizontal line across the center of the graph represents the output at quiescence. For Figure 6.3, the constants are arranged so that the main laser is fully quenched when ω_c is at the peak gain (and $\omega_c=\omega$). The semiconductor gain appears on the right-hand axis with that for QW1 magnified by a factor of 2.6. The interaction of the control beam with the gain medium decreases (quenches) or increases (pumps) the output from the quiescent value depending on the magnitude of the photon energy $\hbar\omega_c$.

As a key point, the gain for the control beam is different from the net gain for the main-laser mode. Although the two beams P and P_c interact with the same isotropic gain medium, they have different boundary conditions and produce different results. The cavity beam P approximates a standing wave so that the Fabry-Perot resonance condition

along with the gain condition $G \cong (\tau_p v \Gamma)^{-1}$ produce a narrow laser linewidth. The control beam, however, interacts with the semiconductor according to the dynamics embodied in the gain G but without the Fabry-Perot effects. As a result, the optical bandwidth of the gain for the control beam is equal to the bandwidth of the material gain and not the width of the laser line.

6.4 Discussion of the Observed Characteristics

The description of the spontaneous emission region (SER) in Figure 2 requires the wavelength dependence of the transfer function. First consider Figure 3 and a non-integrated control source. The optical bandwidth B_Q for the "quench spectrum" and B_P for the "pump spectrum" of the main laser is represented by the portion of the curves in the lower and upper half of the figure. As previously noted, B_Q is identical to the bandwidth of the semiconductor material gain. Generally, larger B_Q should obtain from lasers with larger threshold current I_{thr} (larger η_{thr}) or fewer quantum wells since the peak gain per well must be larger.[16] As a primary result, the effectiveness of spontaneous emission (SE) for quenching the laser is primarily determined by the amount of overlap between the SE and quench spectra. If B_S is the bandwidth for spontaneous emission, then an elementary statistical argument suggests that the approximate maximum total bandwidth B_T for quenching by SE is on the order of $\frac{1}{B_T^2} = \frac{1}{B_Q^2} + \frac{1}{B_S^2}$ (for exact overlap $B_S=B_Q$ and $B_T=B_Q/\sqrt{2}$). Larger B_Q implies greater quenching by spontaneous emission.

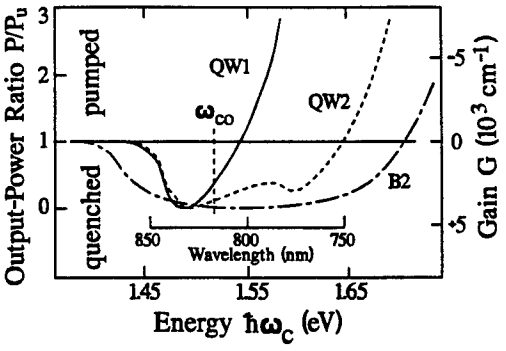


Figure 6.3: Graphical depiction of quenching transfer function for a laser with a single quantum well (QW1 and QW2) and one made from bulk semiconductor (B1) as discussed in the text. QW1 is electrically pumped to a carrier density of 5×10^{18} (cm^{-3}), while QW2 and B2 are pumped to 12.5×10^{18} (the active regions are 80 angstoms thick). The right hand scale shows the gain (after ref. 10) for QW2 and B2 while the QW1 curve is multiplied by a factor of approximately 2.6.

The SER shown in Figure 6.2 for a monolithically integrated control laser is explained by the spectral overlap (discussed above) and by the spatial profile of the spontaneous emission (SE). First, the response $\left|\frac{\partial P}{\partial I_c}\right| \cong 0$ tends to be small since the SE is radiated in all directions and only a small fraction $\beta/2$ (from Eq. 6.9) enters the main laser cavity. The response to the SE from the integrated control laser tends to remain small above threshold (for the control laser) since the rate of SE saturates at a level determined by the threshold carrier density. Second, the sign of the slope $\frac{\partial P}{\partial I_c}$ depends on the relative average wavelengths of the three spectra (SE, quench and pump). The center wavelengths of the three specta (SE, quench and pump) can shift with respect to each other even though the main and control lasers are integrated on the same substrate. The fact that I_{thr} decreases as I_c increases (for $P_c<0.1$) indicates significant overlap between the SE and pump spectra; for this device, electrical cross-talk and heating effects are ruled

out. Contrarily, the SER for the top curve of Figure 6.2 shows a slight "quenching" effect.

The linear region of the transfer curves (Figure 6.2) and the associated linear shifts of the main-laser threshold current are described by Equations 6.10 and 6.11. The last term in Equation 6.10 is linear in P_c and can be interpreted as either a threshold current for the main laser (for L-I curves) or as part of the transfer function. The linear region of the transfer function is due to the linear nature of perturbing matter-light interaction V, the form of Liouville Equation and the fact that the magnitude of the gain is fixed by the condition $G \cong (\tau_p v \Gamma)^{-1}$ at the lasing wavelength. As a result, the data in Figure 2 is fit with the straight line[8,14] $\frac{P}{P_u} = 1 - mb\frac{P_c}{P_u}$ (the same form as Equation 11) where {b=2.1 mA/mW, m=0.42 mW/mA, η_{cm}=2} for Q and {b=0.78, m=0.093, η_{cm} =0.17} for H (assuming $\omega=\omega_c$ and $R_c=R_m \cong 0.33$). As a note, the coupling efficiency η_{cm}, which measures the fraction of the control beam power coupled into the side of the main-cavity, appears to be outside the acceptable range of 0 to 1 for the Q device. We believe that this anomalous value is due to electrical conduction between the lasers through the nonzero resistance of the substrate and bottom electrode.[15] Also note that the optical gain $G_{opt} \equiv \left|\frac{dP}{dP_c}\right|$ is 0.16 and 0.33 for H and Q, respectively.

The Saturation Region (SR) can be characterized by a zero or nonzero asymptote. If the main-laser is fully quenched, then it will only spontaneously emit and the asymptote will be near P=0. Although the SR is not fully addressed in this work, we speculate that this region can be explained by shifts between the main-laser gain curve and the control laser output spectrum and by spatial hole burning in the main laser.

6.5 New Device Applications and Predictions

The transfer curves and the associated bandwidths are particularly important for applications of wavelength converters, FM demodulators and optical amplifiers. Wavelength conversion occurs when an *amplitude modulated* (AM) control beam with optical frequency ω_{co} *amplitude modulates* the main laser operating at optical frequency ω_o (positioned at the peak gain for mirrors without wavelength selectivity). Optical FM to AM demodulation (wavelength demodulation) occurs when an FM signal with frequency ω_c varying about the quiescent value ω_{co} is applied to the side of the main cavity. If the deviation of $\omega_c(t)$ from ω_{co} is small compared with the bandwidth of $G(\omega_c)$ then, by a Taylor expansion of $G(\omega_c)$ in Equation 11, the deviation of the output power can be written as

$$\Delta P = -\frac{\eta_{cm} A}{G(\omega_o)} (\Delta\omega_c) P_c \left[\frac{\partial G(\omega_c)}{\partial \omega_c}\right]_{\omega_c=\omega_{co}} \tag{6.12}$$

where P_c is the optical power of the carrier for the FM signal. Equation 6.12 explicitly shows how the instantaneous amplitude of the AM output beam depends on the instantaneous deviation in frequency ($\Delta\omega_c=\omega_c-\omega_{co}$), the differential frequency gain $\partial G/\partial\omega_c$, and the power of the carrier P_c for the input FM signal. As seen from Figure 6.3, the differential frequency gain $\partial G/\partial\omega_c$ depends on the magnitude of the bias current and the bandwidth for optical quenching. In principle, the conversion efficiency $\frac{\Delta P}{\Delta\omega_c}$ can be dynamically adjusted by applying a voltage to a semiconductor (voltage controlled) absorber section that might form part of the main cavity. As another new result, the optically controlled lasers can exhibit optical gain G_{opt}. Equation 10 predicts

$$G_{opt} = \left|\frac{dP}{dP_c}\right| = \eta_{cm} A \frac{G(\omega_c)}{G(\omega)} \tag{6.13}$$

Optical gain $G_{opt} > 1$ obtains either by (1) positioning ω_c at the peak material gain (refer to Figure 2) and using frequency selective (DBR) mirrors that have a resonant frequency detuned from the peak material gain (i.e., $G(\omega_c)>G(\omega)$) or (2) adjusting the ratio of beam areas or reflectivities in A.

6.6 Summary

In summary, the control mechanisms are identified as gain saturation and gain competition. The wavelength-dependent quenched-laser transfer function has been derived from the Liouville Equation for the density matrix for multiple optical fields. This steady-state transfer function was shown to be essential for understanding and applying the optically-controlled lasers. Additionally, the transfer function provided new information on optical gain, wavelength conversion and demodulation for optically quenched lasers. The derivation has shown that the main lasers can be either a VCSEL or an edge-emitting laser fabricated from a wide variety of materials.

Appendix 1: Review of Integrating Factors

Now for a quick review of integrating factors as a method of solving first order differential equations. Suppose we want to solve the equation

$$\dot{y} - ay = f(t) \tag{A1.1}$$

where y=y(t) and the dot indicates the first derivative with respect to time. Suppose we multiply through by a function $\mu(t)$, the integrating factor,

$$\mu\dot{y} - a\mu y = \mu f(t) \tag{A1.2}$$

with the particular property that the left-hand side is an exact derivative

$$\frac{d}{dt}(\mu y) = \mu\dot{y} - a\mu y \tag{A1.3}$$

Then we could write Equation A1.2 as

$$\frac{d}{dt}(\mu y) = \mu f(t)$$

If the forcing function f(t) starts at t=0, we can integrate both sides of the equation with respect to time to obtain

$$\mu(t)y(t) = \mu(0)y(0) + \int_0^t d\tau\, \mu(\tau)f(\tau)$$

or

$$y(t) = \frac{\mu(0)y(0)}{\mu(t)} + \frac{1}{\mu(t)}\int_0^t d\tau\, \mu(\tau)f(\tau) \tag{A1.4}$$

So once the integrating factor $\mu(t)$ *is known then so the solution is also known even when the exact form of the forcing function has not been specified.* This is the property that makes the integrating factor useful for our puposes in the next topic.

How do we find the integrating factor? Use Equation A1.3 and expand the derivative

$$\frac{d}{dt}(\mu y) = \mu\dot{y} + \dot{\mu}y \tag{A1.5}$$

Combining equations A1.3 and A1.5 we find

$$\mu\dot{y} + \dot{\mu}y = \mu\dot{y} - a\mu y$$

to arrive at

$$\dot{\mu} = -a\mu$$

By separating variables, this simple first order differential equation has the solution

$$\mu(t) = e^{-at}$$

Notice that we did not include the constant of integration since they are unnecessary for integrating factors.

Appendix 2: Brief Review of Special Relativity

The basic mathematical construct in relativity theory is Minkowski space which consists of the set of four-vectors. Four-vectors transform according to the Lorentz transformation which is a rotation in Minkoski space. The four-vector in any moving reference frame can be determined from that in any other frame using the Lorentz transformation. In many cases, it is easiest to calculate quantities in a "rest frame" and then apply the Lorentz transformation to find the corresponding quantities in the moving reference frame. The Lorentz transformation is a result of the experimental fact that the speed of light is a constant regardless of the state of motion of the observer. There is no "stationary" coordinate system in the universe and therefore no *naturally* preferred reference frame. The mathematical formulation of valid physical laws must be independent of any particular reference frame; this means that the equations must be applicable to any coordinate system regardless of its state of uniform motion. Mathematical expressions that are valid for all reference frames are termed "relativistically covariant"; i.e., they retain their form under a Lorentz transformation. As a note, the theory of relativity is becoming increasingly important in a number of areas of engineering. The operation the free electron laser requires the theory as discussed in Yariv's book in Chapter 13.

The identification of the four-vectors is important for physical theory. The list of four-vectors include position-time

$$x_\mu = (x_1, x_2, x_3, x_4) = (\vec{x}, ix_o) = (\vec{x}, ict)$$

(where $i = \sqrt{-1}$), the four-gradient

$$\partial_\mu = \frac{\partial}{\partial x_\mu} = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}, \frac{\partial}{\partial x_4}\right) = \left(\nabla, \frac{1}{ic}\frac{\partial}{\partial t}\right)$$

momentum-energy

$$p_\mu = (c\vec{p}, iE)$$

(where E is the total energy and not just the energy of the rest mass)
the vector-scalar potential

$$A_\mu = (\vec{A}, iA_o)$$

the current-charge density

$$J_\mu = (\vec{J}, ic\rho)$$

If the components of a four-vector are known in one reference frame (i.e., coordinate system) then they are known in any other through the Lorentz transformation. As is conventional, we assume that the motion between two reference frames is along the $z=x_3$ axis and that the unprimed system observes the primed system to be moving along the positive z-direction with speed v. The trick here is to picture the mathematical rotations in Minkowski space rather than the motion between reference systems.

First consider rotations in Euclidean space (basically, the complex "i" makes Minkowki space an Euclidean space). The figure shows a rotation of a 2-D vector $\vec{r}$ by an angle θ which is equivalent to rotating the reference frame by $-\theta$. The rotated vector is related to the original one by the operator $\hat{R}$

$$|r'\rangle = \hat{R}|r\rangle$$

which has the matrix

$$\underline{R} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

The Lorentz transformation rotates the x_3 and x_4 components for motion along the z-direction according to

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \\ x'_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta & -\sin\theta \\ 0 & 0 & \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

where the other components x_1 and x_2 are unaffected by motion along the z-direction. The same transformation holds for all of the different types of four-vectors. The transformation equation can be written in terms of typical parameters by the following definitions

$$\theta = i\alpha \qquad \beta = \frac{v}{c} \qquad \gamma = \frac{1}{\sqrt{1-\beta^2}} \qquad \beta = \tanh(\alpha)$$

where "tanh" is the hyperbolic tangent and the last relation leads to

$$\cosh(\alpha) = \frac{1}{\sqrt{1-\beta^2}} = \gamma \qquad \sinh(\alpha) = \frac{\beta}{\sqrt{1-\beta^2}} = \beta\gamma$$

We have

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \\ ict' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta & -\sin\theta \\ 0 & 0 & \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ ict \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cosh(\alpha) & i\sinh(\alpha) \\ 0 & 0 & -i\sinh(\alpha) & \cosh(\alpha) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ ict \end{pmatrix}$$

or, more simply

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \\ ict' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & i\beta\gamma \\ 0 & 0 & -i\beta\gamma & \gamma \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ ict \end{pmatrix}$$

The discussion above shows that four-vectors transform according to

$$x'_\mu = R_{\mu\nu} x_\mu$$

so that the components in one reference frame can be related to the components in a second one in uniform motion (along z) with respect to the first. The transformation matrix is

$$R_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & i\beta\gamma \\ 0 & 0 & -i\beta\gamma & \gamma \end{pmatrix} \qquad \mu,\nu = 1,2,3,4$$

Rotations in Minkowski space are orthogonal in the sense that $\underline{R}^{-1} = \underline{R}^{T}$ and the length of a four-vector

$$x_\mu x_\mu = \sum_{\mu=1}^{4} x_\mu x_\mu = \vec{x}\cdot\vec{x} - c^2t^2$$

is left invariant under the transformation. Note the convention of an implied sum over repeated indices. The invariance is easy to see using matrix notation

$$x'_\mu\, x'_\mu \equiv \underline{x}'^{T}\,\underline{x}' = (\underline{R}\underline{x})^{T}(\underline{R}\underline{x}) = \underline{x}\underline{R}^{T}\underline{R}\underline{x} = \underline{x}\underline{x} = x_\mu x_\mu = 1$$

The length of a four-vector is therefore a scalar under the Lorentz transformation.

The scalar quality of the length of a four-vector provides the momentum-energy relation. Starting with $p_\mu = (c\vec{p}, iE)$, we find

$$p_\mu p_\mu = c^2\vec{p}^{\,2} - E^2$$

However, in a reference frame at rest with respect to the particle $\vec{p} = 0$ and only the rest mass contributes to the total energy of the particle $E = mc^2$. The length of the energy-momentum four-vector in any reference frame is given by

$$p_\mu p_\mu = c^2\vec{p}^{\,2} - E^2 = -(mc^2)^2$$

where E now consists of the rest mass and the kinetic energy. The previous equation leads to the famous result

$$E = \sqrt{c^2\vec{p}^{\,2} + (mc^2)^2}$$

Tensors

Tensors $F_{\mu\nu}$ transform according to

$$F_{\alpha\beta} = R_{\alpha\mu}R_{\beta\nu}F_{\mu\nu}$$

where repeated indices are summed. Once the components of the tensor $F_{\mu\nu}$ are known in one reference frame, they are known in all others in uniform motion with respect to the first. One expecially nice example concerns the electromagnetic field. We can show that a magnetic field is really an electric field in motion! That is, if we have an electric field due to a stationary point charge in one frame, then in a second frame in uniform motion, an observer will see both electric and magnetic fields! The motion between the two frames has converted a portion of the electric field into a magnetic field.

Appendix 3 shows that the tensor is given by

$$F_{\mu\nu} = \begin{bmatrix} 0 & H_z & -H_y & -icD_x \\ -H_z & 0 & H_x & -icD_y \\ H_y & -H_x & 0 & -icD_z \\ icD_x & icD_y & icD_z & 0 \end{bmatrix} = \begin{bmatrix} 0 & B_z/\mu_o & -B_y/\mu_o & -ic\varepsilon_o E_x \\ -B_z/\mu_o & 0 & B_x/\mu_o & -ic\varepsilon_o E_y \\ B_y/\mu_o & -B_x/\mu_o & 0 & -ic\varepsilon_o E_z \\ ic\varepsilon_o E_x & ic\varepsilon_o E_y & ic\varepsilon_o E_z & 0 \end{bmatrix}$$

The field tensor is usually stated in the Lorentz units (need to check the MKS units above). Suppose in one frame an observer sees an electric field along the x-axis

$$F_{\mu\nu} = \begin{bmatrix} 0 & 0 & 0 & -ic\varepsilon_o E_x \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ ic\varepsilon_o E_x & 0 & 0 & 0 \end{bmatrix}$$

The Lorentz transformation then gives the magnetic field F'_{13} in the moving frame.

$$F'_{13} = R_{1\mu} R_{3\nu} F_{\mu\nu}$$

The only contribution to the summations is from F_{14}. We need

$$R_{11} = 1 \qquad R_{34} = i\beta\gamma$$

to find

$$-\frac{B'_y}{\mu_o} = F'_{13} = R_{11} R_{34} F_{14} = i\beta\gamma\left(-ic\varepsilon_o E_x\right)$$

The (absolute value) magnetic field seen in the moving reference frame is

$$B'_y = \frac{\beta}{c}\gamma E_x$$

Similarly we can show that the electric field in the moving reference frame is

$$E'_x = \gamma E_x$$

In the moving frame, the ratio of the magnetic to electric field is

$$\frac{B'_y}{E'_y} = \frac{\beta}{c}$$

Suppose that the original frame is actually attached to a beam of light so that the second frame move at the velocity of light with respect to the first. In this case,

$$\beta = \frac{v}{c} = 1$$

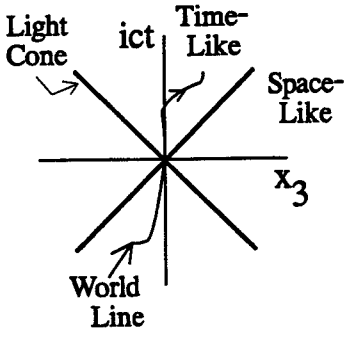
and the ratio of the electric fields is

$$\frac{B'_y}{E'_y} = \frac{1}{c}$$

just as has been found in previous chapters.

Space-Time Plots

A space-time plot is divided into three regions depending on whether or not the origin can be connected to the points by a signal without exceeding the speed of light. A four-vector V_μ is time-like if $\vec{V}^2 < V_4^2$, space-like if $\vec{V}^2 > V_4^2$ and light-like if $\vec{V}^2 = V_4^2$. Points in the space-like region would require signals to propagate faster than the speed of light for the signal to reach them from the origin. A world line is created by a particle as it moves through space-time. A differential element of length along the world line is found from Pythagorus relation



$$(dL)^2 = \sum_\mu (dx_\mu)^2 = (d\vec{x})\cdot(d\vec{x}) - c^2(dt)^2$$

which is independent of coordinate system. The differential "proper time" $d\tau$ is defined to be the differential length of the position four-vector as measured in the reference frame at rest with the particle (i.e., the reference frame travels with the particle). In this case, dx=0 and so $dL = ic\, d\tau$. The time interval $d\tau$ is measured by a clock at rest with the moving particle. Using the fact that the length of the four-vector is invariant under a Lorentz transformation (the length is a scalar), the differential interval in any reference frame has the value

$$(d\tau)^2 = -\frac{1}{c^2}\sum_{\mu}(dx'_{\mu})(dx'_{\mu})$$

which leads to the usual time-dilation formula. The four-velocity is defined to be

$$dv_{\mu} = \frac{dx_{\mu}}{d\tau}$$

which is a valid four-vector since $d\tau$ is a scalar.

Appendix 3: Covariant form of Maxwell's Equations and the Lorentz Gauge

There are a number of Gauge transformations for electromagnetics that includes the Coulomb gauge, Lorentz gauge, temporal gauge and axial gauge. Some of the gauges are apparent only from a covariant formulation; Appendix 1 reviews four-vector notation.

In MKS units, Maxwell's equations in free space can be written in covariant form as

$$\partial_\mu F_{\nu\mu} = J_\nu$$

(summation over μ) where

$$\partial_\mu = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}, \frac{1}{ic}\frac{\partial}{\partial t}\right) \qquad J_\nu = \left(\vec{J}, ic\rho\right)$$

and the electromagnetic field tensor is

$$F_{\mu\nu} = \begin{bmatrix} 0 & H_z & -H_y & -icD_x \\ -H_z & 0 & H_x & -icD_y \\ H_y & -H_x & 0 & -icD_z \\ icD_x & icD_y & icD_z & 0 \end{bmatrix} = \begin{bmatrix} 0 & B_z/\mu_o & -B_y/\mu_o & -ic\varepsilon_o E_x \\ -B_z/\mu_o & 0 & B_x/\mu_o & -ic\varepsilon_o E_y \\ B_y/\mu_o & -B_x/\mu_o & 0 & -ic\varepsilon_o E_z \\ ic\varepsilon_o E_x & ic\varepsilon_o E_y & ic\varepsilon_o E_z & 0 \end{bmatrix}$$

The units need to be checked on the field-tensor (since it is usually given in Lorentz units). Greek subscripts usually refer to all four components (whereas subscripts such as "j" refer to the first three). For free space

Notice that $F_{\mu\nu}$ is antisymmetric $F_{\mu\nu} = -F_{\nu\mu}$

Example: Find $\partial_\mu F_{j\mu} = J_j$

Solution:

$$\partial_\mu F_{1\mu} = \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} - \frac{ic}{ic}\frac{\partial D_x}{\partial t} = J_1$$

The 2nd and 3rd components are similar. The result is

$$\nabla \times \vec{H} = \frac{\partial \vec{D}}{\partial t} + \vec{J}$$

The field tensor can be related to a vector potential by writing

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

where $A_\nu = \left(\vec{A}, iA_o\right)$.

Example: Use the four-vector potential A_μ to find $\nabla \cdot \vec{B}$

Solution: Write

$$\nabla \cdot B = \mu_o\{\partial_1 F_{23} - \partial_2 F_{13} + \partial_3 F_{12}\} = \mu_o\{\partial_1 F_{23} + \partial_2 F_{31} + \partial_3 F_{12}\}$$
$$= \mu_o \partial_1(\partial_2 A_3 - \partial_3 A_2) + \ldots = 0$$

Maxwell's equations as represented by

$$\partial_\nu F_{\mu\nu} = J_\mu$$

can be rewritten by substituting

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

to get

$$\partial_\mu \partial_\nu A_\nu - \partial_\nu \partial_\nu A_\mu = J_\mu$$

The last equation can be rewritten using a 4-D Laplacian operator

$$\Box = \partial_\nu \partial_\nu = \nabla^2 - \frac{1}{c^2}\frac{\partial}{\partial t^2}$$

as

$$\Box A_\mu - \partial_\mu \partial_\nu A_\nu = -J_\mu \qquad \text{(A3.1)}$$

which is similar to the wave equation for the four-potential except for the added term of $-\partial_\mu \partial_\nu A_\nu$

Lorentz Gauge

A new 4-potential A_μ^{new}, in the Lorentz gauge, can be defined that allows the wave equation to be written as

$$\Box A_\mu^{new} = -J_\mu \qquad \text{(A3.2)}$$

where the new 4-potential satisfies the Lorentz condition

$$\partial_\mu A_\mu^{new} = 0$$

The new wave equation in the Lorentz gauge essentially obtains by noting that the physical objects of interest are the electromagnetic fields $F_{\mu\nu}$. The derivatives in

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

provide the freedom to redefine the "old" 4-potential A_μ while maintaining the fields $F_{\mu\nu}$. Define

$$A_\mu^{new} = A_\mu + \partial_\mu \chi \qquad \text{(A3.3)}$$

where $\chi = \chi(\vec{r}, t)$ is a scalar function.
We can see that the "old" and "new" electromagnetic fields are the same

$$F_{\mu\nu}^{new} = \partial_\mu A_\nu^{new} - \partial_\nu A_\mu^{new} = \partial_\mu (A_\nu + \partial_\nu \chi) - \partial_\nu (A_\mu + \partial_\mu \chi) = F_{\mu\nu} + (\partial_\mu \partial_\nu \chi - \partial_\nu \partial_\mu \chi) = F_{\mu\nu}$$

since order of the derivatives can be rearranged in the last set of parenthesis so that it is zero.

The new wave equation (A3.2) can be demonstrated from Equation A3.1 by solving Equation A3.3 for the old potential A_μ

$$A_\mu = A_\mu^{new} - \partial_\mu \chi$$

in the form of

$$\Box A_\mu = \Box A_\mu^{new} - \partial_\mu \Box \chi$$

and substituting into Equation A3.1

$$\Box A_\mu - \partial_\mu \partial_\nu A_\nu = -J_\mu$$

to get

$$\Box A_\mu^{new} - \partial_\mu \Box \chi - \partial_\mu \partial_\nu A_\nu = -J_\mu$$

Now choose the gauge function χ such that

$$\Box \chi = -\partial_\nu A_\nu \qquad \text{(A3.4)}$$

The new wave equation is then

$$\Box A_{\mu}^{new} = -J_{\mu}$$

Finally the Lorentz gauge condition

$$\partial_{\mu} A_{\mu}^{new} = 0$$

comes from Equation A3.3

$$A_{\mu}^{new} = A_{\mu} + \partial_{\mu}\chi$$

as can be seen by taking the 4-divergence of both sides

$$\partial_{\mu} A_{\mu}^{new} = \partial_{\mu} A_{\mu} + \partial_{\mu}\partial_{\mu}\chi = \partial_{\mu} A_{\mu} + \Box\chi = 0$$

where the last step follows by Equation A3.4.

Appendix 4: The Origin of Fock Space

There are many ways to categorize particles. One method divides particles into Fermions and bosons. Fermions are particles that have half-integer spin such as 1/2, 3/2,.... Electrons are Fermions. Bosons are particles that have integer spin such as 0,1,2.... Photons and phonons are Bosons. Spin is pictured as the rotation of a classical object and is related to its angular momentum. Bosons and phonons obey different statistics which leads to different symmetry properties of the wave functions and different types of operator commutation relations. It all stems from Pauli's exclusion principle which says that only one *Fermion* can occupy a particular state at one time. Any number of *bosons* can occupy the same state at the same time. For a system consisting of many Fermions (such as electrons) described by the Fock state $|n_1,n_2,...\rangle$, the integers $n_1,n_2,...$ can only take on the values of 0 or 1. These facts are represented by certain properties of the Fock states and the wavefunctions that represent the system of particles. This section shows how the Fock states are related to the wavefunctions and how the properties of Fermions and Bosons are manifested. The final section contains a comment on the Pauli exclusion principle. (This appendix follows Chapter 1 of Fetter and Walecka with some notation changes).

Section A4.1: Origin of Fock States

Assume a system of N particles. At this point, it doesn't matter whether they are Fermions or bosons. The particles have wave functions that depend on the coordinates x_k and the time. Assume that the Hamiltonian has the form

$$\hat{H} = \sum_k \hat{T}(x_k) + \frac{1}{2}\sum_{\substack{k,j \\ k\neq j}} \hat{V}(x_k,x_j) \tag{A4.1}$$

where the kinetic energy $\hat{T}$ might have the form $\hat{T}_k = \frac{\hat{p}_k^2}{2m_k} = -\frac{\hbar^2}{2m_k}\frac{\partial^2}{\partial x_k^2}$ (but it doesn't need to have this form) and the potential term might have the form of Coulomb interaction

$$\hat{V}(x_k,x_j) \sim \frac{1}{|x_k - x_j|}$$

The summation over the potential terms does not include j=k since that term is a self interaction term and the potential would be infinite. The factor of ½ occurs since $\hat{V}(x_k,x_j) = \hat{V}(x_j,x_k)$ and we don't want to include the same term twice.

The general wavefunction has the form

$$\psi(x_1,x_2,...,x_N,t) = \sum_{E_1,E_2...E_N} C(E_1,E_2...,E_N,t)u_{E_1}(x_1)u_{E_2}(x_2)...u_{E_N}(x_N) \tag{A4.2}$$

and solves the many-body Schrodinger equation

$$\hat{H}\psi = i\hbar\frac{\partial}{\partial t}\psi \tag{A4.3}$$

The basis set $\{u_E(x)\}$ consists of single-body wavefunctions that account for the boundary conditions and the set $\{E\}$ consists of the corresponding energy eigenvalues. Notice, as usual, the basis set is independent of time. The subscripts on "x" and "E" refer to the particle number. For example, an infinitely deep well has energy eigenstates that are sines or cosines with energy eigenvalues given by E_n. If there is only one particle, the general wavefunction must be some combination of the basis set such as

$$\psi(x,t) = \sum_E C(E,t)\,u_E(x)$$

as is usual. For each particle "k", there is a Hilbert space with a basis set $\{u_E(t)\}$ and so one might think that the general wavefunction for all N particles should be

$$\psi(x_1,\ldots,x_N,t) = \sum_{E_1,\ldots,E_N} C_1(E_1,t)u_{E_1}(x_1)\,C_2(E_2,t)u_{E_2}(x_2)\ldots C_N(E_N,t)u_{E_N}(x_N)$$

but this is not the most general form for a vector in a direct product of Hilbert spaces. The correct form is given in Equation A4.2. In practice, all of the quantum numbers in the E_k (such as energy, angular momentum etc) should be included; however, the discussion follows the simplest path. Keep in mind that the *position* in the arguments of $\psi(\ldots,x_i,\ldots,x_j,\ldots)$ or in $C(E_1,E_2,\ldots,E_N,t)$ refer to a particular particle and not necessarily the x_i. In principle, the set of wavefunctions should be superscripted with an "(i)" to indicate the particle number so that the general wavefunction would read

$$\psi(x_1,x_2,\ldots,x_N,t) = \sum_{E_1,E_2\ldots E_N} C(E_1,E_2\ldots,E_N,t)u^{(1)}_{E_1}(x_1)u^{(2)}_{E_2}(x_2)\ldots u^{(N)}_{E_N}(x_N)$$

since each particle "(i)" occupies its own Hilbert space spanned by the set $\{u^{(i)}_{E_i}\}$ where E_i takes on the range of eigenvalues. However, the "i" on the E_i consistently indicates the Hilbert space number.

To see how the Fock states arise, consider first the symmetry property of the wave functions which is related to the statistical properties of the particles. Afterwards consider a new counting scheme that redefines the expansion coefficients

$$C(E_1,E_2,\ldots,E_N,t)$$

and results in the Fock states.

Start with the observation that bosons and fermions obey different symmetry properties when two particles are interchanged; i.e., the position coordinates of the particles are interchanged. We require

$$\psi(\ldots,x_i,\ldots,x_j,\ldots) = \pm\psi(\ldots,x_j,\ldots,x_i,\ldots) \qquad \text{(A4.4)}$$

where "+" refers to bosons and "-" refers to Fermions. The choice leads to the particle statistics and Pauli's exclusion principle. It can be shown that interchanging the particle coordinates in Equation A4.4 is equivalent to interchanging the energy labels in "C" according to

$$C(\ldots,E_i,\ldots,E_j,\ldots,t) = \pm C(\ldots,E_j,\ldots,E_i,\ldots,t) \qquad \text{(A4.5)}$$

where again "+" refers to bosons and "-" refers to Fermions. To see this using only a two-particle system, start with Equation A4.4 and substitute Equation A4.2 for ψ on both sides to obtain

$$\sum_{E_1,E_2} C(E_1,E_2,t)\,u_{E_1}(x_1)u_{E_2}(x_2) = \pm\sum_{E_1,E_2} C(E_1,E_2,t)\,u_{E_1}(x_2)u_{E_2}(x_1)$$

On the right-hand side, interchange the dummy indices E_1, E_2 to obtain

$$\sum_{E_1,E_2} C(E_1,E_2,t)\, u_{E_1}(x_1) u_{E_2}(x_2) = \pm \sum_{E_1,E_2} C(E_2,E_1,t)\, u_{E_2}(x_2) u_{E_1}(x_1)$$

Compare both sides to obtain the results in Equation A4.5.

Topic A4.1.1: Bosons

Now use the symmetry of the coefficients to show the origin of the Fock state for bosons. First redefine the coefficients as follows. The eigenvalues for all Hilbert spaces take on identical ranges of values. Here, consider the range $\{E_i\} = \{1,2,3,\ldots\}$ for every space "i". The energies E_i with the lowest values (i.e., "1") are moved to the left in the coefficients $C(E_1,E_2,\ldots,t)$ (which can be accomplished by using the symmetry property in Equation A4.5). Let "n_1" be the number of particles with energy "1" and so on. Then one can write

$$C(E_1,E_2,\ldots,t) = C\left(\begin{array}{c} E_a,E_b,\ldots E_c,E_d,\ldots,E_e,\ldots, \\ \leftarrow n_1 \rightarrow \leftarrow n_2 \rightarrow \end{array}\right)$$

Define a new coefficient $\bar{C}$ with an argument that has positions corresponding to energy rather than particle number. We write

$$C(E_1,E_2,\ldots,t) = \bar{C}(n_1,n_2,\ldots,n_\infty,t)$$

where obviously

$$N = \sum_{i=1}^{n_\infty} n_i$$

is the total number of particles.

Example: Suppose that there are three particles and five energy states

$$\{E_i : i = 1,2,3\} = \{1,2,3,4,5\}$$

then, for example, the coefficient C(2,1,1) can be written

$$C(E_1 = 2, E_2 = 1, E_3 = 1) = C(2,1,1) = C(1,1,2) = \bar{C}(n_1 = 2, n_2 = 1, n_3 = 0,\ldots) = \bar{C}(2,1,0,0,\ldots)$$

Now we can rewrite the general wave function in Equation A4.2

$$\psi(x_1,x_2,\ldots,x_N,t) = \sum_{E_1,E_2\ldots E_N} C(E_1,E_2\ldots,E_N,t) u_{E_1}(x_1) u_{E_2}(x_2)\ldots u_{E_N}(x_N)$$

as

$$\psi(x_1,x_2,\ldots,x_N,t) = \sum_{n_1,n_2,\ldots n_\infty} \sum_{\substack{E_1,E_2\ldots E_N \\ (n_1,n_2,\ldots n_\infty)}} \bar{C}(n_1,n_2\ldots,n_\infty,t) u_{E_1}(x_1) u_{E_2}(x_2)\ldots u_{E_N}(x_N) \qquad \text{(A4.5)}$$

where the notation "$(n_1,n_2,\ldots,n_\infty)$" at the bottom of the second summation symbol means to hold the number of particles n_1, n_2,... constant while performing the summation. The following example shows the meaning of the restricted summation and indicates that the summations in the previous equations are just an alternate method of adding over all energies.

Example: Consider the case of three particles and five energy levels as in the previous example. Assume the restriction that $n_1=2$ and $n_2=1$ and $n_i=0$ for i=3,4,5. The allowed configurations are

$$E_1=1 \quad E_2=1 \quad E_3=2$$
$$E_1=1 \quad E_2=2 \quad E_3=1$$
$$E_1=2 \quad E_2=1 \quad E_3=1$$

Therefore the restricted summation can be evaluated

$$\sum_{\substack{E_1,E_2\ldots E_N \\ (n_1,n_2,\ldots n_\infty)}} C(E_1,E_2,E_3) = C(1,1,2)+C(1,2,1)+C(2,1,1) = 3\,C(1,1,2) = 3\bar{C}(2,1,0,0,\ldots)$$

The restricted summation adds over all the energy while keeping a constant number of particles with a particular energy.

The Fock states obtain from Equation A4.5, which is

$$\psi(x_1,x_2,\ldots,x_N,t) = \sum_{n_1,n_2,\ldots n_\infty} \sum_{\substack{E_1,E_2\ldots E_N \\ (n_1,n_2,\ldots n_\infty)}} \bar{C}(n_1,n_2\ldots,n_\infty,t)\,u_{E_1}(x_1)\,u_{E_2}(x_2)\ldots u_{E_N}(x_N)$$

by defining new expansion coefficients

$$\beta(n_1,n_2\ldots,n_\infty,t) = \left(\frac{N!}{n_1!n_2!\ldots n_\infty!}\right)^{1/2} \bar{C}(n_1,n_2\ldots,n_\infty,t) \qquad \text{(A4.6)}$$

and an alternate set of basis vectors according to the prescription

$$\phi_{n_1,n_2,\ldots n_\infty}(x_1,x_2,\ldots,x_N) = \left(\frac{n_1!n_2!\ldots n_\infty!}{N!}\right)^{1/2} \sum_{\substack{E_1,E_2\ldots E_N \\ (n_1,n_2,\ldots n_\infty)}} u_{E_1}(x_1)\,u_{E_2}(x_2)\ldots u_{E_N}(x_N) \qquad \text{(A4.7)}$$

The new basis vector $\phi_{n_1,n_2,\ldots n_\infty}$ is the Fock state

$$|n_1,n_2,\ldots,n_\infty\rangle = \left(\frac{n_1!n_2!\ldots n_\infty!}{N!}\right)^{1/2} \sum_{\substack{E_1,E_2\ldots E_N \\ (n_1,n_2,\ldots n_\infty)}} |u_{E_1}\rangle\,|u_{E_2}\rangle\ldots|u_{E_N}\rangle$$

projected into coordinate space. Each Fock state for different n_i is a different basis vector as seen in the previous section. The general wave function now has the form

$$\psi(x_1,x_2,\ldots,x_N,t) = \sum_{n_1,n_2,\ldots n_\infty} \beta(n_1,n_2\ldots,n_\infty,t)\,\phi_{n_1,n_2,\ldots n_\infty}(x_1,x_2,\ldots,x_N) \qquad \text{(A4.8)}$$

More detailed treatments actually show that the Fock states are correctly normalize

$$\langle \phi_{n_1,n_2,\ldots n_\infty}(x_1,x_2,\ldots,x_N) | \phi_{m_1,m_2,\ldots m_\infty}(x_1,x_2,\ldots,x_N)\rangle = \delta_{n_1m_1}\delta_{n_2m_2}\cdots$$

by actually calculating the inner product.

Topic A4.1.2: Fermions

It is possible to use the same reasoning for the Fermion Case. The antisymmetry of the wavefunction under interchange of coordinates in Equations A4.4 and A4.5

$$\psi(\ldots,x_i,\ldots,x_j,\ldots) = -\psi(\ldots,x_j,\ldots,x_i,\ldots) \qquad \text{(A4.4)}$$
$$C(\ldots,E_i,\ldots,E_j,\ldots,t) = -C(\ldots,E_j,\ldots,E_i,\ldots,t) \qquad \text{(A4.5)}$$

The Fermion Fock states obtain from Equation A4.5

$$\psi(x_1,x_2,\ldots,x_N,t)=\sum_{n_1,n_2,\ldots n_\infty}\sum_{\substack{E_1,E_2\ldots E_N\\(n_1,n_2,\ldots n_\infty)}}\bar{C}(n_1,n_2\ldots,n_\infty,t)u_{E_1}(x_1)u_{E_2}(x_2)\ldots u_{E_N}(x_N)$$

by defining new expansion coefficients

$$\beta(n_1,n_2\ldots,n_\infty,t)=\left(\frac{N!}{n_1!n_2!\ldots n_\infty!}\right)^{1/2}\bar{C}(n_1,n_2\ldots,n_\infty,t) \qquad \text{(A4.9)}$$

and an alternate set of basis vectors using the determinant (sometimes called the Slater determinant)

$$\phi_{n_1,n_2,\ldots n_\infty}(x_1,x_2,\ldots,x_N)=\left(\frac{n_1!n_2!\ldots n_\infty!}{N!}\right)^{1/2}\begin{vmatrix}u_{E_1}(x_1) & \cdots & u_{E_1}(x_N)\\ \vdots & & \vdots\\ u_{E_N}(x_1) & \cdots & u_{E_N}(x_N)\end{vmatrix} \qquad \text{(A4.10)}$$

The last equation is the Fock state $|n_1,n_2,\ldots,n_\infty\rangle$ projected into coordinate space. Each Fock state for different n_i is a different basis vector as seen in the previous topic. The general wave function now has the form

$$\psi(x_1,x_2,\ldots,x_N,t)=\sum_{n_1,n_2,\ldots n_\infty}\beta(n_1,n_2\ldots,n_\infty,t)\phi_{n_1,n_2,\ldots n_\infty}(x_1,x_2,\ldots,x_N) \qquad \text{(A4.11)}$$

More detailed treatments actually show that the Fock states are correctly normalize

$$\langle\phi_{n_1,n_2,\ldots n_\infty}(x_1,x_2,\ldots,x_N)|\phi_{m_1,m_2,\ldots m_\infty}(x_1,x_2,\ldots,x_N)\rangle=\delta_{n_1m_1}\delta_{n_2m_2}\cdots$$

by actually calculating the inner product.

Example: What is the coordinate representation of the Fock state for a two electron system with only two energy levels?

Solution: Using Equation A4.10, one finds

$$\phi_{n_1,n_2,\ldots n_\infty}(x_1,x_2)=\left(\frac{n_1!n_2!\ldots n_\infty!}{2!}\right)^{1/2}\begin{vmatrix}u_{E_1}(x_1) & u_{E_1}(x_2)\\ u_{E_2}(x_1) & u_{E_2}(x_2)\end{vmatrix}$$

$$=\left(\frac{n_1!n_2!\ldots n_\infty!}{2!}\right)^{1/2}\left(u_{E_1}(x_1)u_{E_2}(x_2)-u_{E_2}(x_1)u_{E_1}(x_2)\right)$$

where n_i =0,1 due to the Pauli exclusion principle. Therefore,

$$\phi_{n_1,n_2,\ldots n_\infty}(x_1,x_2)=\frac{1}{\sqrt{2}}\left(u_{E_1}(x_1)u_{E_2}(x_2)-u_{E_2}(x_1)u_{E_1}(x_2)\right)$$

Notice that E_1=E_2 is forbidden since the wavefunction ϕ would be zero.

Section A4.2: The origin of Boson Creation and Annihilation Operators

Returning to the original statement of the problem for an N-particle system without regard to the distinction between Fermions and Bosons, recall the Hamiltonian and general wavefunction.

$$\hat{H} = \sum_{k} \hat{T}(x_k) + \frac{1}{2} \sum_{\substack{k,j \\ k \neq j}} \hat{V}(x_k, x_j) \tag{A4.12}$$

The general wavefunction satisfying the many body Schrodinger Equation

$$\hat{H}\psi = i\hbar \frac{\partial}{\partial t}\psi \tag{A4.13}$$

has the form

$$\psi(x_1, x_2, \ldots, x_N, t) = \sum_{W_1, W_2 \ldots W_N} C(W_1, W_2 \ldots, W_N, t)\, u_{W_1}(x_1)\, u_{W_2}(x_2) \ldots u_{W_N}(x_N) \tag{A4.14}$$

where the notation has been changed for later convenience. The set $\{W_i\}$ denotes the energy eigenvalues for the particle #i. Substituting Equation A4.14 into A4.13, provides

$$\sum_{W_1, W_2 \ldots W_N} \left[\frac{\partial}{\partial t} C(W_1, \ldots, W_N, t) \right] u_{W_1}(x_1) \ldots u_{W_N}(x_N) =$$

$$= \sum_{W_1, \ldots, W_N} C(W_1, \ldots, W_N, t) \left[\sum_{k} \hat{T}(x_k) + \frac{1}{2} \sum_{\substack{k,j \\ k \neq j}} \hat{V}(x_k, x_j) \right] u_{W_1}(x_1) \ldots u_{W_N}(x_N)$$

Factor out the two summations on the right-hand side, multiply from the right by the *operator*

$$\int dx_1 \ldots dx_N \; u^*_{E_1}(x_1) \ldots u^*_{E_N}(x_N)$$

(where E_1, E_2,... are now specific energy values) to find

$$\sum_{W_1, W_2 \ldots W_N} \left[\frac{\partial}{\partial t} C(W_1, \ldots, W_N, t) \right] \int dx_1 \ldots dx_N \; u^*_{E_1}(x_1) \ldots u^*_{E_N}(x_N)\, u_{W_1}(x_1) \ldots u_{W_N}(x_N) =$$

$$= \sum_{W_1, \ldots, W_N} C(W_1, \ldots, W_N, t) \int dx_1 \ldots dx_N \; u^*_{E_1}(x_1) \ldots u^*_{E_N}(x_N) \left[\sum_{k} \hat{T}(x_k) \right] u_{W_1}(x_1) \ldots u_{W_N}(x_N)$$

$$+ \sum_{W_1, \ldots, W_N} C(W_1, \ldots, W_N, t) \int dx_1 \ldots dx_N \; u^*_{E_1}(x_1) \ldots u^*_{E_N}(x_N) \left[\frac{1}{2} \sum_{\substack{k,j \\ k \neq j}} \hat{V}(x_k, x_j) \right] u_{W_1}(x_1) \ldots u_{W_N}(x_N)$$

The functions $u_{E_j}(x_j)$ are a particular choice of the basis functions so that the orthonormality relations

$$\int dx_j u^*_E(x_j) u_W(x_j) = \delta_{E,W}$$

can be used to simplify the equations (notice both functions in the integral have the same coordinates). The result is

$$\frac{\partial}{\partial t}C(E_1,\ldots,E_N,t) =$$

$$= \sum_k \sum_{W_k} C(E_1,\ldots E_{k-1},W_k,E_{k+1},\ldots t) \int dx_k\, u^*_{E_k}(x_k)\,\hat{T}(x_k)\,u_{W_k}(x_k)$$

$$+ \sum_{\substack{k,j \\ k\neq j}} \sum_{W_k W_j} C(E_1,\ldots W_j,E_{j+1},\ldots W_k,E_{k+1}\ldots,t) \int dx_j dx_k\, u^*_{E_j}(x_j) u^*_{E_k}(x_k) \frac{1}{2}\hat{V}(x_k,x_j)\,u_{W_j}(x_j) u_{W_k}(x_k)$$

Once again restrict the argument to bosons. Consider the coefficient $C(E_1,\ldots E_{k-1},E_k,E_{k+1},\ldots t)$ with the corresponding number coefficient given by

$$\bar{C}(n_1,n_2,\ldots,n_{E_k},\ldots t) = C(E_1,\ldots,E_k,\ldots t)$$

where n_{E_k} means the number of particles with the energy E_k. The coefficient $C(E_1,\ldots E_{k-1},W_k,E_{k+1},\ldots t)$ changes the energy E_k of particle #k to the new energy W_k. There is one less particle with energy E_k and one more with W_k. Therefore,

$$C(E_1,\ldots E_{k-1},W_k,E_{k+1},\ldots t) = \bar{C}(n_1,\ldots,n_{E_k-1},\ldots,n_{W_k}+1,\ldots t)$$

This can be incorporated in the kinetic energy term

$$ke = \sum_k \sum_{W_k} C(E_1,\ldots E_{k-1},W_k,E_{k+1},\ldots,t) \int dx_k\, u^*_{E_k}(x_k)\,\hat{T}(x_k)\,u_{W_k}(x_k)$$

by noting a general sum of the form

$$\sum_k f(E_k) = f(a)+f(b)+\ldots$$

where the symbols a,b,c... represent one of the possible energy values E. The terms in the sum can be grouped according to the different energy values

$$\sum_k f(k) = \underset{\leftarrow n_1 \rightarrow}{f(a)+f(j)+\ldots} + \underset{\leftarrow n_2 \rightarrow}{f(d)+\ldots+\ldots.} = \sum_E n_E f(E)$$

Therefore, the kinetic energy term becomes

$$ke = \sum_k \sum_{W_k} C(E_1,\ldots E_{k-1},W_k,E_{k+1},\ldots,t) \int dx_k\, u^*_{E_k}(x_k)\,\hat{T}(x_k)\,u_{W_k}(x_k)$$

$$= \sum_E n_E \bar{C}(n_1,n_2,\ldots,n_E-1,\ldots,n_W+1,\ldots,t)\langle E|\hat{T}|W\rangle$$

Let i,j now represent the energy values, we can write

$$ke = \sum_{EW} n_E \bar{C}(n_1,n_2,\ldots,n_E-1,\ldots,n_W+1,\ldots,t)\langle E|\hat{T}|W\rangle$$

$$= \sum_{ij} n_i \langle i|\hat{T}|j\rangle \bar{C}(n_1,n_2,\ldots,n_i-1,\ldots,n_j+1,\ldots,t)$$

Fetter and Walecka also evaluate the potential energy term. When the two results are combine with the coefficients from Equation A4.6

$$\beta(n_1,n_2\ldots,n_\infty,t) = \left(\frac{N!}{n_1!n_2!\ldots n_\infty!}\right)^{1/2} \bar{C}(n_1,n_2\ldots,n_\infty,t)$$

they end up with a messy looking equation.

$$i\hbar\frac{\partial}{\partial t}\beta(n_1,n_2\ldots,n_\infty,t)=\sum_i\langle i|\hat{T}|i\rangle\, n_i\,\beta(n_1,\ldots,n_i,\ldots n_\infty,t)+$$

$$+\sum_{\substack{ij\\ i\neq j}}\langle i|\hat{T}|j\rangle\sqrt{n_i}\sqrt{n_j+1}\,\beta(n_1,n_2,\ldots,n_i-1,\ldots,n_j+1,\ldots,t)+$$

$$+\sum_{i\neq j\neq k\neq m}\langle ij|\hat{V}|km\rangle\frac{1}{2}\sqrt{n_i}\sqrt{n_j}\sqrt{n_k+1}\sqrt{n_m+1}\,\beta(\ldots,n_i-1,\ldots,n_j-1,\ldots,n_k+1,\ldots,n_m+1,\ldots,t)$$

$$+\sum_{i=j\neq k\neq m}\langle ii|\hat{V}|km\rangle\frac{1}{2}\sqrt{n_i}\sqrt{n_i-1}\sqrt{n_k+1}\sqrt{n_m+1}\,\beta(\ldots,n_i-2,\ldots,n_k+1,\ldots,n_m+1,\ldots,t)$$

$$+\text{ETC}$$

There is one of these long equations for each set of occupation numbers $n_1,n_2,\ldots$

We can now proceed as follows. Using the Schrodinger equation

$$i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle=\hat{H}|\Psi(t)\rangle \tag{A4.15}$$

where

$$\psi(x_1,x_2,\ldots,x_N,t)=\sum_{n_1,n_2,\ldots n_\infty}\beta(n_1,n_2\ldots,n_\infty,t)\,\phi_{n_1,n_2,\ldots n_\infty}(x_1,x_2,\ldots,x_N)$$

or

$$|\psi(t)\rangle=\sum_{n_1,n_2,\ldots n_\infty}\beta(n_1,n_2\ldots,n_\infty,t)|n_1,n_2,\ldots n_\infty\rangle \tag{A4.16}$$

By substituting A4.16 in A4.15 and working with the Hamiltonian,

$$i\hbar\sum_{n_1,n_2,\ldots n_\infty}\frac{\partial\beta(n_1,n_2\ldots,n_\infty,t)}{\partial t}|n_1,n_2,\ldots n_\infty\rangle=\hat{H}|\Psi(t)\rangle \tag{A4.17}$$

The expression for the derivative of β (long equation above) can be substituted into A4.17 to yield an alternate expression for $\hat{H}$. The second kinetic energy term

$$i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle=\ldots+\sum_{n_1,n_2,\ldots,n_\infty}\sum_{\substack{ij\\ i\neq j}}\langle i|\hat{T}|j\rangle\beta(\ldots,n_i-1,\ldots,n_j+1,\ldots,t)\sqrt{n_i}\sqrt{n_j+1}|n_1,\ldots,n_\infty\rangle+\ldots \tag{A4.18}$$

Notice that the square roots and the Fock state are almost the form required for creation and annihilation operators. Redefine the dummy indices according to

$$n_i-1\rightarrow n_i \qquad n_j+1\rightarrow n_j$$

to get

$$i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle=\ldots+\sum_{n_1,n_2,\ldots,n_\infty}\sum_{\substack{ij\\ i\neq j}}\langle i|\hat{T}|j\rangle\beta(\ldots,n_i,\ldots,n_j,\ldots,t)\sqrt{n_i+1}\sqrt{n_j}|\ldots n_i+1,\ldots n_j-1\ldots\rangle+\ldots$$

Now we can substitute the creation and annihilation operators to get

$$i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle=\ldots+\sum_{n_1,n_2,\ldots,n_\infty}\sum_{\substack{ij\\ i\neq j}}\langle i|\hat{T}|j\rangle\beta(\ldots,n_i,\ldots,n_j,\ldots,t)\hat{b}_i^+\hat{b}_j|\ldots n_i,\ldots n_j-1\ldots\rangle+\ldots$$

All of the terms in the expansion A4.18 can be rewritten in terms of the creation and annihilation operators. The result is

$$\hat{H} = \sum_{i,j} b_i^+ \langle i|\hat{T}|j\rangle b_j + \frac{1}{2}\sum_{ijkm} b_i^+ b_j^+ \langle ij|\hat{V}|km\rangle b_k b_m$$

Section A4.3: The Pauli Exclusion Principle.

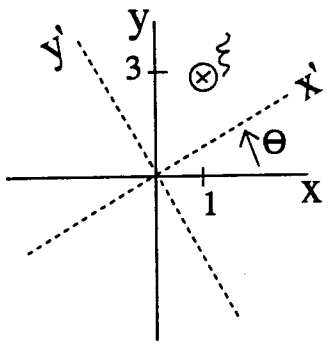
The Pauli Exclusion Principle applies to Fermions. The functions $\{u_1^{(i)}, u_2^{(i)}, \ldots\}$ (one set for each particle "i") span a Hilbert space. Independent particles occupy their own Hilbert space and there is no interaction term connecting them. Consider the case for three particles for simplicity. One can form a direct product space spanned by the basis set $\{u_a^{(1)} u_b^{(2)} u_c^{(3)}\}$ where "a,b,c" all range over the eigenstates of the Hamiltonian for the space. However, the Slater determinant defines the basis set as

$$\psi = C \begin{vmatrix} u_a(x_1) & u_b(x_1) & u_c(x_1) \\ u_a(x_2) & u_b(x_2) & u_b(x_2) \\ u_a(x_3) & u_b(x_3) & u_c(x_3) \end{vmatrix}$$

which is zero if, for example, a=b. This means that certain basis vectors in the direct product space are not legitimate basis vectors for Fermions. For example, the basis vector $u_1(x_1)u_1(x_2)u_3(x_2)$ is not included for the Fermions.

Appendix 5: Review of Functions with Rotated Coordinates

If a person knows a function $f(x,y)$ in one set of coordinates (x,y) then what is the function $f'(x',y')$ for coordinates (x',y') that are rotated through an angle θ with respect to the first set (x,y).



Consider a point in space ξ as indicated in the picture. The single point can be described by the primed or unprimed coordinate system. The key fact is that the equations linking the two coordinate systems describe the single point ξ. The equations for coordinate rotations are

$$\underline{r}' = \underline{R}\,\underline{r} \tag{A5.1.1}$$

where

$$\underline{r}' = \begin{pmatrix} x' \\ y' \end{pmatrix} \qquad \underline{R} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \qquad \underline{r} = \begin{pmatrix} x \\ y \end{pmatrix} \tag{A5.1.2}$$

and $\underline{r}'$ and $\underline{r}$ represent the single point ξ.

A value "z" associated with the point ξ is the same value regardless of the reference frame. Therefore, we require

$$z = f'(x',y') = f(x,y) \tag{A5.1.3}$$

since (x',y') and (x,y) specify the same point ξ. We can write the last equation using Equation A5.1.1 as

$$f'(x',y') = f(x,y) = f(\underline{R}^{-1}\underline{r}') \tag{A5.1.4}$$

where for the depicted 2-D rotation

$$\underline{R} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Example: Suppose the value associated with the point $\underline{r} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ is 10 (i.e., f(1,3)=10) what is $f'(x'=3, y'=-1)$ for $\theta = \pi/2$?

Solution: Using Equation A5.1.4, we find

$$f'(3,-1) = f\left[\underline{R}^{-1}\underline{r}'\right] = f\left[\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} 3 \\ -1 \end{pmatrix}\right] = f\left[\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 3 \\ -1 \end{pmatrix}\right] = f(1,3) = 10$$

Appendix 6: Density Operator and the Boltzmann Distribution

For simplicity, suppose the reservoir consists of a single harmonic oscillator with the Hamiltonian given by

$$\hat{H}_r = \hbar\omega\left(\hat{R}^+\hat{R}^- + \frac{1}{2}\right) = \hbar\omega\left(\hat{N} + \frac{1}{2}\right)$$

where $\hat{R}^+, \hat{R}^-, \hat{N}$ are the raising, lowering and number operators. The density operator $\hat{\rho}_r$ is defined through a Boltzmann distribution

$$\hat{\rho}_r = \frac{1}{Z}\exp\left(-\frac{\hat{H}_r}{k_B T}\right)$$

where Z is the normalization (partition function)

$$Z = \mathrm{Tr}_r\left\{\exp\left(-\frac{\hat{H}_r}{k_B T}\right)\right\}$$

Consider the average of an operator $\hat{O}$

$$\left\langle\hat{O}\right\rangle = \mathrm{Tr}\left(\hat{\rho}_r\hat{O}\right) = \sum_n \langle n|\hat{\rho}_r\hat{O}|n\rangle = \sum_{n,m}\langle n|\hat{\rho}_r|m\rangle\langle m|\hat{O}|n\rangle$$

where the closure relation for the *energy* basis set $\{|n\rangle\}$ is inserted between the two operators. The energy eigenstates are chosen for the basis since the density operator is diagonal in that basis set. First, evaluate the matrix elements of the density operator.

$$\langle n|\hat{\rho}_r|m\rangle = \frac{1}{Z}\langle n|\exp\left(-\frac{\hat{H}_r}{k_B T}\right)|m\rangle = \frac{1}{Z}\langle n|m\rangle\exp\left(-\frac{E_m}{k_B T}\right)$$

where 1/Z is removed from the inner product because it is a c-number and where the last term obtains by operating with the Hamilton on the ket $|m\rangle$. Using the orthogonality of the basis provides

$$\langle n|\hat{\rho}_r|m\rangle = \frac{\delta_{nm}}{Z}\exp\left(-\frac{E_n}{k_B T}\right)$$

and the average of the operator is

$$\left\langle\hat{O}\right\rangle = \mathrm{Tr}\left(\hat{\rho}_r\hat{O}\right) = \sum_n \frac{1}{Z}\exp\left(-\frac{E_n}{k_B T}\right)O_{nn}$$

where O_{nn} are the diagnonal elements in the matrix of $\hat{O}$. The partition function is evaluated similarly. The expectation value of the operator $\hat{O}$ shows that the density operator for the reservoir gives rise to the Boltzmann probability distribution. The energy levels E_n are expected to be populated according to the thermal distribution.

Misc. Applications

+G.S. Agarwal, J. Mod. Optics 34, 909 (1987), "Wigner-function description of quantum noise in interferometers." Discusses optical transfer function for the Mach-Zehnder interferometer using quantum states of light. Primarily theory.
+C.R.Gilson, S.M.Barnett, S. Stenholm, J. Mod. Opt. 34, 949 (1987), "Squeezed quantum fluctuations and noise limits in amplifiers and attenuators." Importance of low-noise linear operators and how squeezing improve the performance. Primarily theory.
+G. Rempe, H. Walther, Phys. Rev A42, 1650 (1990), "Sub-Poissonian atomic statistics in a microMASER," Paper discusses microwave application of squeezed light for single-atom maser. Also see Phys. Rev. Lett. 58, 353 (1987), Phys. Rev. Lett. 64, 2783 (1990)
+H. Ritsch, P. Zoller, Opt. Comm. 64, 523 (1987), "Absorption spectrum of a two level system in a squeezed vacuum." Discusses pump-probe technique for measuring response of two level atom in squeezed vacuum.
+E. Yablonovitch, Phys. Rev. Lett. 58, 2059 (1987), "Inhibited spontaneous emission in solid state physics and electronics." Classic paper describing how to increase transistor, photodetector efficiency and laser efficiency by decreasing the spontaneous emission. This paper started the field of photonic crystals which can be used to decrease the spontaneous emission. See also Yablonovitch et.al, Phys. Rev. Lett. 61, 2546 (1988), Phys. Rev. Lett. 63, 1950 (1989),

Detectors

+G.L. Abbas, V.W.S Chan, T.K. Yee, Opt. Lett. 8, 419 (1983) "Local-oscillator excess-noise suppression for homodyne and heterodyne detection". Discusses homodyne and heterodyne optical detection systems for communications and sensor systems. Calculates S/N ratio includes quantum and classical effects.
+M.J.Collett, R. Loudon,C.W.Gardiner, J. Mod. Opt. 34,881 (1987), "Quantum theory of optical homodyne and heterodyne detection". Paper discusses the theory for detection of arbitrary quantum electomagnetic states.
+R.H. Dicke, "The Measurement of Thermal Radiation at Microwave Frequencies," Rev. Sci. Instr. 17, 268 (1946). First homodyne detection system for radar.
+M.G. Raymer, J. Cooper, H.J. Carmichael, M. Beck, D.T. Smithey, "Ultrafast measurement of optical-field statistics by dc-balanced homodyne detection" J. Opt. Soc. Am. B 12, 1801 (1995). Good review of homodyne detection.
+T.Waite, A Balanced Mixer for Optical Heterodyning: The Magic T Optical Mixer, Proc. IEEE, Page 334, February 1966.

Laser Applications

+G.S. Agarwal, S. Dutta Gupta, Phys. Rev. A39, 2961, (1989), "Dynamical interaction of an atomic oscillator with squeezed radiation inside a cavity" Discusses rate equations for the interaction between squeezed electromagnetic states and atoms inside/outside a cavity. Theory

+C.W.Gardiner, Phys. Rev. Lett. 56, 1917 (1986), "Inhibition of Atomic Phase Decays by Squeezed Light: A Direct Effect of Squeezing." Squeezed vacuums decrease the spontaneous emission rate. Theory. See also Phys. Rev A40, 3796 (1989).

+J.Gea-Banacloche, Phys. Rev. Lett. 59, 543 (1987), "Squeezing of Spontaneous Emission in a Laser." Squeezed vacuums reduce spontaneous emission noise and reduce the laser line width. Theory

+E.F.Goodbar, R.J. Ram, R. Nagarajan, L.A.Coldren, J.E. Bowers, Appl. Phys. Lett. 66, 3419 (1995) "Intensity noise and facet correlation in Fabry-Perot laser diodes with low facet reflectivity." Discusses sources of noise in semiconductor lasers with cleaved facets. Experiment

+C. Harder, J. Katz, J. Shacham, A. Yariv, IEEE J. Quantum Electr. QE-18, 333 (1982), "Noise Equivalent Circuit of a Semiconductor Laser Diode." Rate equations for semiconductor laser including Langevin noise sources. Engineering model.

+S. Inoue, S. Lathi, Y. Yamamoto, "Longitudinal-mode-partition noise and amplitude squeezing in semiconductor lasers" J. Opt. Soc. Am. B 14, 2761 (1997). Facet reflectivity is not the only factor determining degree of amplitude squeezing.

+J. Kitching, A. Yariv, Y. Shevy, Phys. Rev. Lett. 74, 3372 (1995), "Room temperature generation of amplitude squeezed light from a semiconductor laser with weak optical feedback." The paper discusses how to produce amplitude squeezed light using a semiconductor laser and a current-source. Low noise detection. Experiment.

+J. Kitching, R. Boyd, A. Yariv, Y. Shevy, Opt. Lett. 19, 1331 (1994), "Amplitude noise reduction in semiconductor lasers with weak, dispersive optical feedback." Experimental work with descriptive equations.

+S. Machida, Y. Yamamoto, Y. Itaya, Phys. Rev. Lett. 58, 1000 (1987), "Observation of Amplitude Squeezing in a constant-current-driven semiconductor laser." Thermal noise is reduced by using a large biasing resistor as a current source. Uses a DFB laser. Experiment. See also Phys. Rev A35, 5114 (1987) which discusses the mechanism in more detail with an Historical perspective. See also Opt. Lett. 8 177 (1983) first describes the correct noise source in homodyne and heterodyne detection systems.

+M.A.M. Marte, H. Ritsch, D.F. Walls, Phys. Rev. A38, 3577 (1988), "Squeezed Reservoir Lasers." See especially the references for using squeezed light to stabilize the output of a laser.

+R. Loe-Mie et.al., Optics Communic. 139, 55 (1997), "Nearly all-optical frequency-stabilization of a laser diode on the 120 kHz intercombination line of Ba." Shows how to stabilize a semiconductor laser to 100 kHz.

+Machida, S., Yamamoto, Y., "Observation of amplitude squeezing from semiconductor lasers by balanced direct detectors with a delay line" Opt. Lett. 14, 1045 (1989).

+W. H. Richardson, R. M. Shelby, Phys. Rev. Lett. 64, 400 (1990), "Nonclassical light from a semiconductor laser operating at 4K". See also Richardson et.al., Phys. Rev. Lett. 66, 2867 (1991), "Squeezed photon-number noise and sub-poisson electrical partition noise in a semiconductor laser." Good discussion of experiment and laser construction factors that affect laser noise. They squeeze upto 96% of the noise out of the light.
+Y. Yamamoto, N. Imoto, IEEE J. Quant. Electr. QE-22, 2032 (1986), "Internal and External Field Fluctuations of a Laser Oscillator: Part I-quantum Mechanical Langevin Treatment." Discusses distinction between emitted laser beam and intra-cavity beam.

Basic Physics/Engineering and Mathematics
+H.P. Yuen, Phys. Rev A13, 2226 (1976), "Two-photon coherent states of the radiation field." This is paper that started the notion of squeezed states of the electromagnetic field.
+P. Adam, J. Janszky, Phys. Rev. A50, 3548 (1994), "Comment on 'squeezing and antisqueezing for a harmonic oscillator having a sudden change of mass." The paper shows that sudden changes of the mass of a harmonic oscillator can result in a squeezed wavefunction.
+S.M. Barnett, D.T. Pegg, J. Mod. Opt. 36, 7 (1989), "On the Hermitian Optical Phase Operator" Defines non-commuting phase and number operators for the electromagnetic field.
+G. Barton, Proc. of Roy. Soc. Lond. A320, 251 (1970), Proc. of Roy. Soc. Lond. A410, 175 (1987). Quantum mechanical shifts in frequency of light of emitting light and inhibition of spontaneous emission.
+H.J.Carmichael, A.S.Lane, D.F.Walls, J. Mod. Opt. 34, 821 (1987), "Resonance fluorescence in a squeezed vacuum". Discusses resonant fluorescence from an atom in a coherent optical field when the atom is damped by a squeezed-vacuum reservoir.
+H.J.Carmichael, A.S.Lane, D.F.Walls, Phys. Rev. Lett. 58, 2539 (1987), "Resonance Fluorescence from an Atom in a Squeezed Vacuum." Similar to above.
+C.M. Caves, Phys. Rev. D23, 1693 (1981), "Quantum noise in an interferometer." Brief review of important issues on squeezed states. Application to an interferometer for measuring distances. Theory.
+L. Davidovich, J.M.Raimond,M.Brune,S.Haroche, Phys. Rev. A36, 3771 (1987), "Quantum theory of a two-photon micromaser." Develops theory for wide tuning range and possible generation of squeezed light.
+R. Graham, J. Mod. Opt. 34, 873 (1987), "Squeezing and frequency changes in harmonic oscillators" Discusses method of squeezing matter wavefunctions.
+G. Hu, E. Irwin, P.K. Aravind, J. Mod. Opt. 42, 375 (1995), "The damped-vacuum-field Jaynes-Cummings model."
+Jaynes, E. T., Cummings, F.W., "Comparison of Quantum and Semiclassical Radiation Theories with Application to the Beam Maser" Proc. IEEE, January 1963 page 89. Discusses the Jaynes-Cummings' model.
+R. Matloob, R. Loudon, Phys. Rev. A53, 4567 (1996), "Electromagnetic field quantization in absorbing dielectrics II.

General Purpose

+J.E. Carroll, Optica Acta 33, 909 (1986), "A proposal for quiet photon streams from injection lasers." Gives suggestion on how to make quiet semiconductor lasers using built-in feedback mechanisms.
+B. Dahmani, L. Hollberg, R. Drullinger, Opt. Lett. 12, 876 (1987), "Frequency stabilization of semiconductor lasers by resonant optical feedback." Shows how linewidths can be reduced to 20 kHz using an external Fabre-Perot cavity.
+M.S. Feld, K. An, Scientific American page 57 (July 1998), "The Single-Atom Laser" Nice introduction to single atom lasers and the related quantum mechanical effects.
+M. Hawton, Phys. Rev. A50, 1057 (1994), "One-photon operators and the role of vacuum fluctuations in the Casimir force." Interesting discussion of how photon creation near a charged particle can eliminate the need for vacuum fluctuations throughout space.
+I.D. Johnston, Am. J. Phys. 64, 245 (1996), "Photon states made easy: A computational approach to quantum radiation theory." Discusses using packaged software for quantum optic calculations.

Reviews

+L. Davidovich, Rev. Mod. Phys. 68, 127 (1996), "Sub-Poissonian processes in quantum optics."
+R. Loudon, P.L.Knight, J. Mod. Opt. 34, 709 (1987), "Squeezed Light"
+C.H. Henry, R.F. Kazarinov, Rev. Mod. Phys. 68, 801 (1996), "Quantum Noise in Photonics"
+ Proceedings of the IEEE April 1999, "Quantum Devices and Applications" See also the April 1997 issue for review of nanometer scale components.
+B.E.A. Saleh, M.C. Teich, Proc. IEEE 80, 451 (1992), "Information Transmission with Photon-Number-Squeezed Light." Excellent easy-to-read review article on squeezed states.

*U.S. GOVERNMENT PRINTING OFFICE: 1999-510-079-81205